

28128

STIC-Biotech/ChemLib

From: Turner, Sharon
Sent: Wednesday, April 26, 2000 9:55 AM
To: STIC-Biotech/ChemLib
Subject: 09080127

Please search CDB and Interference files

AAs SEQ ID NO2 residues 16-496

NAs SEQ ID NO1 residues 46-1488

AAs SEQ ID NO2 with backtranslates against NA database with oligomer search

NAs SEQ ID NO1 with oligomer search

Sharon L. Turner, Ph.D.

CM1-8A12 GAU 1644

(703) 308-0056

Point of Contact:
Beverly Shears
Technical Info. Specialist
CM1 12C14 Tel: 308-4994

1/26
2/26
3/26
4/26
5/26
6/26
7/26
8/26
9/26
10/26
11/26
12/26

1/26
2/26
3/26
4/26
5/26
6/26
7/26
8/26
9/26
10/26
11/26
12/26

new client

Call 1-800-767-3369

Access DB# _____

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

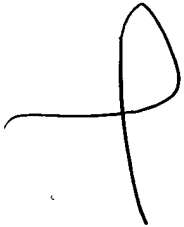
 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*



STAFF USE ONLY

Searcher: Beverly C. 4994
 Searcher Phone #: _____
 Searcher Location: _____
 Date Searcher Picked Up: _____
 Date Completed: 05-11-00
 Searcher Prep & Review Time: _____
 Clerical Prep Time: _____
 Online Time: _____

Type of Search

NA Sequence (#) _____
 AA Sequence (#) _____
 Structure (#) _____
 Bibliographic _____
 Litigation _____
 Fulltext _____
 Patent Family _____
 Other _____

Vendors and cost where applicable

STN _____
 Dialog _____
 Questel/Orbit _____
 Dr.Link _____
 Lexis/Nexis _____
 Sequence Systems _____
 WWW/Internet _____
 Other (specify) CGN

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2000, 19:06:00 ; Search time 32.29 Seconds

(without alignments)
352.834 Million cell updates/sec

Title: US-09-080-127-2_COPY_16_496

Perfect score: 2500
Sequence: 1 GRALVSPDFEPDIEDL.....APKTHAVSGSGCMHSQVRA 481

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 2366106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2500	100.0	496	W89614	Aspergillus oryzae
2	2500	100.0	496	W89597	Aspergillus oryzae
3	2500	100.0	496	W89586	Aspergillus oryzae
4	782.5	31.3	537	W89588	Saccharomyces cere
5	782.5	31.3	537	W89588	Saccharomyces cere
6	668	26.7	500	Y04867	Mycobacterium spec
7	668	26.7	527	Y04868	Secreted protein e
8	229.5	9.2	472	W35604	Human secreted pro
9	222.5	8.9	472	W35604	Human secreted pro
10	188.5	7.5	806	R27481	Prostate-specific
11	161	6.4	751	W31524	Prostate-specific
12	157.5	6.3	393	R39416	Aminopeptidase pre
13	154	6.2	707	W47155	Extracellular doma
14	154	6.2	750	R50597	Prostate-specific
15	154	6.2	750	W02234	Prostate-specific
16	151	6.0	1959	R10562	Mutant protease (d
17	151	6.0	1962	R10560	Mutant protease (K
18	151	6.0	1962	R10561	Mutant protease (N
19	151	6.0	1962	R10557	Mutant protease (A
20	151	6.0	1962	R10558	Mutant protease (A
21	151	6.0	1962	R10559	Mutant protease (A
22	151	6.0	1962	R10563	Mutant protease (K
23	151	6.0	1962	R10564	Mutant protease (d
24	151	6.0	1962	R10565	Mutant protease (d
25	149	6.0	1150	W22471	Streptococcal C4a
26	149	6.0	1150	W22470	Streptococcal C4a
27	147	5.9	1560	P94145	S. cremoris protei
28	145	5.8	1164	W22469	Streptococcal C4a
29	140.5	5.6	622	W22212	Rat transferrin re
30	134.5	5.4	377	W05389	Aminopeptidase. En
31	131.5	5.3	345	W18879	Leucine aminopepti
32	130	5.2	760	W22213	Human transferrin
33	122	4.9	1992	W04505	Moraxella 200 kDa
34	114	4.6	546	W38234	VEGF/FGF fusion p

35	112	4.5	1026	1	R97374	Phage T4 tail fibr
36	109	4.4	1026	1	W92358	Bacteriophage t4 t
37	109	4.4	1026	1	W94678	Bacteriophage T4 t
38	108.5	4.3	447	1	W61228	Streptococcus pneu
39	107	4.3	1262	1	W13505	B. bronchiseptica
40	107	4.3	1644	1	W13504	B. bronchiseptica
41	105.5	4.2	551	1	W18780	Corrected Bacillus
42	105.5	4.2	922	1	W37372	Trimmed enzyme pro
43	102.5	4.1	1018	1	P82115	Fibronectin bindin
44	102.5	4.1	1027	1	W89806	Staphylococcus aur
45	102	4.1	648	1	W80638	S. pneumoniae prot

ALIGNMENTS

RESULT	1	ALIGNMENTS
W89614	188963 standard; Protein; 496 AA.	
ID	W89614;	
AC	W89614;	
DE	17-MAR-1999 (first entry)	
DT	Aspergillus oryzae dipeptidyl aminopeptidase #2.	
DR	Prolyl dipeptidyl aminopeptidase; protein hydrolysate; dough;	
KW	flavour enhancer; palatability; mouthfeel; aroma; crust colour;	
KW	baking; animal feed additive; hydrolysis.	
OS	Aspergillus oryzae.	
PN	W09851803-A1.	
PD	19-NOV-1998.	
PF	12-MAY-1998; U09629.	
PR	20-OCT-1997; US-062892.	
PR	16-MAY-1997; US-857884.	
PA	(NOVO) NOVO NORDISK BIOTECH INC.	
PI	Blinkovsky A, Brown K, Byun T, Klotz A, Rey WM;	
DR	WPI; 99-045232/04.	
DR	N-PSDB: X00070.	
PT	New dipeptidyl aminopeptidase from Aspergillus oryzae - used to	
PT	produce protein hydrolysates enriched in particular amino acids,	
PT	useful as flavour enhancers, e.g. in doughs	
PS	Claim 1; Fig 1; 77pp; English.	
CC	The present sequence represents dipeptidyl aminopeptidase (DPAP) from	
CC	Aspergillus oryzae. DPAP acts synergistically with an aminopeptidase	
CC	(AP) to hydrolyse polypeptides, producing protein hydrolysate (PH),	
CC	useful in foods as flavour enhancer, e.g. in baked goods, enriched in:	
CC	(a) Ala, Arg, Asp, Gly and/or Val, or (b), if the substrate has been	
CC	deaminated, in Glu (free and/or peptide bound), in which case products	
CC	are useful as animal feed additives. DPAP can also be used in flavour-	
CC	improving compositions (optionally containing AP) and in dough pre-mixes,	
CC	also for deactivating enzymes and for converting precursors to mature	
CC	proteins. DPAP increases the level of hydrolysis of proteins and thus	
CC	of flavour development, and a mixture with AP may hydrolyse tripeptides	
CC	that are resistant to either enzyme used alone. PH have improved	
CC	solubility, emulsifying and foaming properties, and products containing	
CC	them have better flavour, palatability, mouthfeel, aroma and crust	
CC	colour.	
SO	Sequence 496 AA;	
QY	1 GRALVSPDFEPDIEDLLEGSQQLDFAYAEERRRVEGKAHDTVNTLYEELKKTG 60	
QY	100.0%; Score 2500; DB 1; Length 496;	
QY	Best Local Similarity 100.0%; Pred. No. 1.4e-210;	
QY	Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	16 GRALVSPDFEPDIEDLLEGSQQLDFAYAEERRRVEGKAHDTVNTLYEELKKTG 75	
QY	1 YDYDYKOPVHLMNSADOTLKVGDEIEAKMTSPSEVETADAVVKNLCSSADYPSD 120	
QY	61 YDYDYKOPVHLMNSADOTLKVGDEIEAKMTSPSEVETADAVVKNLCSSADYPSD 135	
QY	76 YDYDYKOPVHLMNSADOTLKVGDEIEAKMTSPSEVETADAVVKNLCSSADYPSD 135	
QY	121 VEGVALIKRGCEPFGDKSVLAARAKAAASIVYNNVAGSMAGTGAASDGPYSAIVGI 180	
QY	136 VEGVALIKRGCEPFGDKSVLAARAKAAASIVYNNVAGSMAGTGAASDGPYSAIVGI 195	

QY 181 SLEDGKRLKLAAGSVYDVLWDVSKQENRTYNNVAQTGSGPNNVVALGCHTDSVEAG 240
 DB 196 SLEDGKRLKLAAGSVYDVLWDVSKQENRTYNNVAQTGSGPNNVVALGCHTDSVEAG 255
 QY 241 PGINDGSGIISNLVIAKALTOYSVKNNAVRFLEMTAEFGILGSNNYVSHLNATELNKIR 300
 DB 256 PGINDGSGIISNLVIAKALTOYSVKNNAVRFLEMTAEFGILGSNNYVSHLNATELNKIR 315
 QY 301 LYLNFMIMASPNYALMIYDGDGSAFNQSGPAGSAQIEKLFEDYYSIDLPHIPTQFDGRS 360
 DB 316 LYLNFMIMASPNYALMIYDGDGSAFNQSGPAGSAQIEKLFEDYYSIDLPHIPTQFDGRS 375
 QY 361 DYEAFLINGIPSGGLFTGAGIMSEENASRWGQAGVADANYHAAGDNMTNLNHEAFELI 420
 DB 376 DYEAFLINGIPSGGLFTGAGIMSEENASRWGQAGVADANYHAAGDNMTNLNHEAFELI 435
 QY 421 NSKATFAVATYANDLSSIPKRNNTSSLRARATMRPFGRAPKTAHAYSGSCWHSQVE 480
 DB 436 NSKATFAVATYANDLSSIPKRNNTSSLRARATMRPFGRAPKTAHAYSGSCWHSQVE 495
 QY 481 A 481
 DB 496 A 496

RESULT 2
 ID W89597 standard; Protein: 496 AA.
 AC W89597;
 DT 17-MAR-1999 (first entry)
 DE Aspergillus oryzae aminopeptidase II.
 KW Aminopeptidase; protein hydrolysate; baking; dough; flavour enhancer;
 KM mouthfeel; palatability; aroma; hydrolysis; animal feed additive.
 OS Aspergillus oryzae.
 PN WO9851804-A1.
 PD 19-NOV-1998.
 PE 15-MAY-1998: U09940.
 PR 20-OCT-1997: US-062893.
 PR 16-MAY-1997: US-857886.
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 PI Blinovsky A, Brown K, Byun T, Golightly E, Kofod LV;
 DR WPI: 99-045233/04.
 DR N-PSDB: X00018.
 PT New aminopeptidase from Aspergillus oryzae - used to produce protein
 hydrolysates enriched in particular amino acids, used as flavour
 enhancers, e.g. in doughs
 PS Claim 1; Fig 1; 92pp: English.
 CC The present sequence represents aminopeptidase (AP) II from Aspergillus
 oryzae. AP is used in combination with an endopeptidase (EP) to
 hydrolyse polypeptides, producing protein hydrolysate (PH), useful in
 foods as flavour enhancer, e.g. in baked goods, enriched in: (a) leu,
 gly (especially), glu, ser, asp, asn, pro, cys, ala and/or gln, or (b),
 if the substrate has been deamidated, in glu (free and/or peptide bound),
 in which case products are useful as animal feed additives. AP can be
 used in flavour-improving compositions (optionally containing EP) and in
 dough pre-mixes. Also AP can be used for deactivating enzymes and for
 converting precursors to mature proteins. AP increases the level of
 hydrolysis of proteins and thus of flavour development. PH have improved
 CC solubility, emulsifying and foaming properties, and products containing
 CC them have better flavour, palatability and aroma.
 SQ Sequence 496 AA;

Query Match 100.0%; Score 2500; DB 1; Length 496;
 Best Local Similarity 100.0%; Pred. No. 1.4e-210; Indels 0; Gaps 0;
 Matches 481; Conservative 0; Mismatches 0;

QY 1 GRALVSPDEPPEIOLLEGGSOLEDFAAYPERNRVGGKAHDPTVYLYEELKKTG 60
 DB 16 GRALVSPDEPPEIOLLEGGSOLEDFAAYPERNRVGGKAHDPTVYLYEELKKTG 75
 QY 61 YYDVYKQPOVHLMNSADQTLKVGDEIEAKTMTYSPSVEVTADVAVYKNLGCSEADYPSD 120

DB 76 YYDVYKQPOVHLMNSADQTLKVGDEIEAKTMTYSPSVEVTADVAVYKNLGCSEADYPSD 135
 QY 121 VEGKVALIRGECPEFGDKSVLAAKAAASIVYNNVAGSMAGTLCGAOSDKGPYSAIVGI 180
 DB 136 VEGKVALIRGECPEFGDKSVLAAKAAASIVYNNVAGSMAGTLCGAOSDKGPYSAIVGI 195
 QY 181 SLEDGKRLKLAAGSVYDVLWDVSKQENRTYNNVAQTGSGDPPNNVVALGCHTDSVEAG 240
 DB 196 SLEDGKRLKLAAGSVYDVLWDVSKQENRTYNNVAQTGSGDPPNNVVALGCHTDSVEAG 255
 QY 241 PGINDGSGIISNLVIAKALTOYSVKNNAVRFLEMTAEFGILGSNNYVSHLNATELNKIR 300
 DB 256 PGINDGSGIISNLVIAKALTOYSVKNNAVRFLEMTAEFGILGSNNYVSHLNATELNKIR 315
 QY 301 LYLNFMIMASPNYALMIYDGDGSAFNQSGPAGSAQIEKLFEDYYSIDLPHIPTQFDGRS 360
 DB 316 LYLNFMIMASPNYALMIYDGDGSAFNQSGPAGSAQIEKLFEDYYSIDLPHIPTQFDGRS 375
 QY 361 DYEAFLINGIPSGGLFTGAGIMSEENASRWGQAGVADANYHAAGDNMTNLNHEAFELI 420
 DB 376 DYEAFLINGIPSGGLFTGAGIMSEENASRWGQAGVADANYHAAGDNMTNLNHEAFELI 435
 QY 421 NSKATFAVATYANDLSSIPKRNNTSSLRARATMRPFGRAPKTAHAYSGSCWHSQVE 480
 DB 436 NSKATFAVATYANDLSSIPKRNNTSSLRARATMRPFGRAPKTAHAYSGSCWHSQVE 495
 QY 481 A 481
 DB 496 A 496

RESULT 3
 ID W89586 standard; Protein: 496 AA.
 AC W89586;
 DT 17-MAR-1999 (first entry)
 DE Aspergillus oryzae aminopeptidase II.
 KW Aminopeptidase; protein hydrolysate; glycine releasing; protease;
 KM proteinaceous material; flavour; food; baking; animal feed additive;
 KM palatability; hydrolysis; solubility; emulsifying; foaming; aroma;
 KM mouthfeel; crust colour.
 OS Aspergillus oryzae.
 PN WO9851163-A2.
 PD 19-NOV-1998.
 PE 15-MAY-1998: U09998.
 PR 16-DEC-1997: US-069719.
 PR 16-MAY-1997: US-857886.
 PR 20-OCT-1997: US-062893.
 PR 16-DEC-1997: DK-001465.
 PA (ASAHI) ASAHI CHEM IND CO LTD.
 PA (NOVO) NOVO NORDISK AS.
 PI Blinovsky A, Brown K, Byun T, Fujii M, Golightly E,
 PI Kofod LV, Marumotac, Mathiasen TE;
 DR WPI: 99-045177/04.
 DR N-PSDB: V82521.
 PT Production of protein hydrolysate - using protease and enzyme that
 PT releases glycine, useful as flavour improvers in foods and animal
 PT feed additives
 PS Claim 4; Fig 1; 84pp: English.
 CC A method has been developed for the production of protein hydrolysates
 CC (PH) comprising reacting a protein with: (i) at least one polypeptide
 CC with Gly-releasing activity; and (ii) at least one other protease so
 CC that the amount of Gly produced is greater than when the protease is
 CC used alone. PH are used to improve flavour of foods (e.g. baked goods)
 CC and as animal feed additives. PH are preferably also enriched in glu
 CC (free and/or peptide bound), so have improved flavour and palatability.
 CC Addition of a polypeptide with Gly-releasing activity increases the
 CC degree of hydrolysis (or reduces the amount of enzyme needed) and
 CC hydrolysates have better solubility, and emulsifying and foaming
 CC properties. Baked goods containing them have improved aroma, mouthfeel
 CC and crust colour. The present sequence represents Aspergillus oryzae
 CC aminopeptidase II, which is used in the method of the invention.

Sequence 496 AA;

Query Match	100.0%;	Score 2500;	DB 1;	Length 496;
Best Local Similarity	100.0%;	Pred. No. 1.4e-210;		
Matches 481;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	GRVSPDEPEDEDIOTLEDLLEGGSOQLDPFAVPERNRVVGKGAHDOTVATLYKEELTKKG	60
Db	16	GRALVSPPEEFEDTQTEBTEBGGSOQLDPFAVPERNRVVGKGAHDOTVATLYKEELTKKG	75
QY	61	YYDYKPOQVHLMNSADOTLKVGDEIEAKTMTYSPSEVYTAJVAVVKNL:CEADYPSD	120
Db	76	YYDYKPOQVHLMNSADOTLKVGDEIEAKTMTYSPSEVYTAJVAVVKNL:CEADYPSD	135
QY	121	VEGRVALIKRECEFGPGRKSVYLAARAKAAASIVYNNVAGSMAAGTLGAAGSPKGYSAIVGI	180
Db	136	VEGRVALIKRECEFGPGRKSVYLAARAKAAASIVYNNVAGSMAAGTLGAAGSPKGYSAIVGI	195
QY	181	SLEBGOKLIKIAEAGSVSYDLMWDKSKOENRTTYNVVAQTKGQDPNNVVAL3GHTDSYEA6	240
Db	196	SLEBGOKLIKIAEAGSVSYDLMWDKSKOENRTTYNVVAQTKGQDPNNVVAL3GHTDSYEA6	255
QY	241	PGINDDSGGIISNVIYAKALQYQSVKNAVRLPMTAEFGLGSSNTYVSHLNTLNLKIR	300
Db	256	PGINDDSGGIISNVIYAKALQYQSVKNAVRLPMTAEFGLGSSNTYVSHLNTLNLKIR	315
QY	301	LYLNFMDIASPNVYALMTIYDGDGSFNOGSPGASQOIKLEFEDYDSDTLHFHTOPDGRS	360
Db	316	LYLNFMDIASPNVYALMTIYDGDGSFNOGSPGASQOIKLEFEDYDSDTLHFHTOPDGRS	375
QY	361	DYEAFLINGLPSGGLFTGABGIMSEENASRWGQAGVAYDANTYHAAGDNNTNLNHEAFLI	420
Db	376	DYEAFLINGLPSGGLFTGABGIMSEENASRWGQAGVAYDANTYHAAGDNNTNLNHEAFLI	435
QY	421	NSKATAFNAVATYANDLISIPRKNTTSSLHRRATMRPFGRAPKTAHVH:GSCGWSHOYE	480
Db	436	NSKATAFNAVATYANDLISIPRKNTTSSLHRRATMRPFGRAPKTAHVH:GSCGWSHOYE	495
QY	481	A 481	
Db	496	A 496	

RESULT 4

W89598

W89598 standard; Protein; 537 AA.

AC W89598;

DT 17-MAR-1999 (first entry)

DE *Saccharomyces cerevisiae* aminopeptidase II Y.

KW Aminopeptidase; protein hydrolysis; baking; flavour enhancer; mouthfeel; palatability; aroma; hydrolysis; animal feed additive.

OS *Saccharomyces cerevisiae*.

PN W03851804-A1.

PD 19-NOV-1998.

PF 15-MAY-1998; U09940.

PR 20-OCT-1997; US-062893.

PR 16-MAY-1997; US-857886.

FA (NOVO) NOVO NORDISK BIOTECH INC.

P1 Blinkovsky A, Brown K, Ryan T, Golightly E, Kofoed LV;

DR WPI: 99-045233/04.

PT New aminopeptidase from *Aspergillus oryzae* - used to produce protein hydrolysates enriched in particular amino acids, used as flavour enhancers, e.g. in doughs

PS Example 8; Page 74-75; 92pp. English.

CC The present sequence represents aminopeptidase (AP) II Y from *Saccharomyces cerevisiae*. AP is used in combination with an endopeptidase (EP) to hydrolyse polypeptides, producing protein hydrolysate (PH), CC useful in foods as flavour enhancer, e.g. in baked goods, enriched in: CC (a) Leu, Gly (especially), Glu, Ser, Asp, Asn, Pro, Cys, Ala and/or Gln, CC or (b), if the substrate has been deamidated, in Glu (free and/or peptide CC bound), in which case products are useful as animal feed additives. AP CC can be used in flavour-improving compositions (optionally containing EP)

CC and in dough pre-mixes. Also AP can be used for deactivating enzymes and
CC for converting precursors to mature proteins. AP increases the level of
CC hydrolysis of proteins and thus of flavour development. PH have improved
CC solubility, emulsifying and foaming properties, and products containing
CC them have better flavour, palatability and aroma.
90 Sequence 537 AA;

Query Match	31.3%;	Score 782.5;	DB 1;	Length 537;
Best Local Similarity	-40.6%;	Pred. NO. 2.8e-60;		
Matches 187;	Conservative 74;	Mismatches 161;	Indels 39;	Gaps 13

Qy	5	VSPEDEFDIOLEBULLBGSQOUEBFA-----YAPERKNVFPGBKADDTIVNLT---	YEE 55
Db	70	VESELEOORIKYVDLNLATNMTADLYRLAYN SPDDYGH- -TRVIGSGHNKTMEITLNVEDD	127
Qy	56	LKTGYDYVYKQPOVHLMTSMNADQTLKAGDEIE-- -AKTMY- -SPSVE- -VTADVAVAKN	109
Db	128	MO-- -DYIVVSLQ- -BEFALSGKIIFSNJSDAETGKSFANTAPALSPVVDGFGVGLVEIPN	184
Qy	110	LGCEBADIYPSDV-----EGKVALIKRCECPBGDKSVLAAKAKAAALSVYNNVAGSNAATL	164
Db	185	LGCEKRYTASVYPRNHNEKQJALITERKCEPBGDSNLAGKFGFVAVIYDNEPKSKGELH	244
Qy	165	GAOSDKGPFYSATIGISLEDOCKLI- -KLAEGASVYDVLWDSKQENRTYVYNAOATGGD	223
Db	245	GTJGEPYKHYIATVGVPRKYGKAILANIALNIDYSLFLFAMDYSVEFKTONTIADTYHGD	304
Qy	224	PNNVVALGHTDSVEAGPGINDDGGIISMLVIAKALTOY SVKNVAFLEFWTAEEFGLG	283
Db	305	PDNIYALGAHSDSVEEGBGINDDGGTISLNLVAKQJLTHKRNKXVFPFAMMAAEEBGLG	364
Qy	284	SNYYSHNAELNKKIRLYLNFEDMIASPNPALMYDDDSGAFNOSGAGSAQIEKLEFDY	343
Db	365	SNFYAYNTFKKENKIRFEMDYDMASPNYELEYD-- -ANNKENKSGSEELKNLTVDY	420
Qy	344	YDSIDLPHIPEQFGSRSDYEAFILNGIIPSGGLFAGEGIMSEENASRWCQAGAYADANT	403
Db	421	YKAHLNLTIVPEFGSRSDYVEFINNGIIPAGIAGAE-----KNNVNN-----GVLDRCY	471
Qy	404	HAAGDMNTNNEHAFLLNSKATAPAVATYANDLSSIRKNT	444
Db	472	HOLDVDVSNLSMDAFINTNTKILASVATYADSDSEFGPKRET	512

RESULT 5
ID W89568 standard; Protein; 537 AA.
W89568
DT 17-MAR-1999 (first entry)
DE Saccharomyces cerevisiae aminopeptidase II y.
KW Aminopeptidase; protein hydrolysate; glycine releasing; protease;
KW proteaceous material; flavour; food; baking; animal feed additive
KW palatability; hydrolysis; solubility; emulsifying; foaming; aroma
KW multifeed; crust colour.
OS Saccharomyces cerevisiae.
PN W09851163-A2.
PN 19-NOV-1998.
PF 15-MAY-1998; U09398.
PR 16-DEC-1997; US-069719.
PR 16-MAY-1997; US-857886.
PR 20-OCT-1997; US-062893.
PR 16-DEC-1997; DK-001465.
PA (ASAHI) ASahi CHEM IND CO LTD.
PA (NOVO) NOVO NORDISK BIOTECH INC.
PA (NOVO) NOVO-NORDISK AS.
PI Blinovsky A, Brown K, Byun T, Fujii M, Golightly E,
PI Kofoed LV, Marumotac, Mathiansen TE;
PI MPI:99-045177/04.
PT Production of protein hydrolysate - using protease and enzyme thut
PT releases glycine, useful as flavour improves in foods and animal
PT feed additives
PS Example 8; Page 59-60; 84pp; English.

CC A method has been developed for the production of protein hydrolysates
 CC (PH) comprising reacting a protein with: (i) at least one polypeptide
 CC with gly-releasing activity; and (ii) at least one other protease so
 CC that the amount of gly produced is greater than when the protease is
 CC used alone. PH are used to improve flavour of foods (e.g. baked goods)
 CC and as animal feed additives. PH are preferably also enriched in Glu
 CC (free and/or peptide bound), so have improved flavour and palatability.
 CC Addition of a polypeptide with gly-releasing activity increases the
 CC degree of hydrolysis (or reduces the amount of enzyme needed) and
 CC hydrolysates have better solubility, and emulsifying and foaming
 CC properties. Baked goods containing them have improved aroma, mouthfeel
 CC and crust colour. The present sequence represents *Saccharomyces*
 CC cerevisiae aminopeptidase II Y from the present invention.
 CC Sequence 537 AA;

Query Match 31.3%; Score 782.5; DB 1; Length 537;
 Best Local Similarity 40.6%; Pred. No. 2.8e-60;
 Matches 187; Conservative 74; Mismatches 161; Indels 39; Gaps 13;

QY 5 VSPDEFEDIDLEDLESQLEDEP-----YAPERNRFGGAHDDYNYL---YEE 55
 DB 70 VESKLQKIVDDNATAMDLRYRANTSTPDYGP--TRVIGSGHKMTMEYILNVFDD 127
 QY 56 LKTGYDYVYKQPOVHLSNADQTLKVGDEIE--AKMTY--SPSYE-YTADYAVYKN 109
 DB 128 MQ--DIYVVSIG-EFEALSGKIISFNLSDAETGKSFANTFALSPYDGVGKIVELPN 184
 QY 110 LGCEADYPSDV-----EGKVALIRGCEPFGDKSVLAAKAKAAAIYNNVAGSMAGTL 164
 DB 185 LGCEKEDYAVPPRHNKQJALIERGKCPFGDKSNLKGFGFTAVYIDNEPKSKELH 244
 QY 165 GAOSDKGPAISAIYGISLEDDOKLT-KLAEGSVSVLDLVWSKQENRTYVNVVQTKGCD 223
 DB 245 GTLGEPTHTYATYGVPRKVGKKLANIANDISLYPAMSYEFTTQIINADTKRGD 304
 QY 224 PNNVVALGHTDSVEAGPINDSGSISINLYAKALTYSVKNAVRELFWTAEEFGLG 283
 DB 305 PDNIYALGAHSDSYVEGPGINDSGTISLNVAKQLTHFINNNVRAMWAABEEGLG 364
 QY 284 SNYVSHNATELNKIRLYLFNFMASPNYALMTYDGGSAFNOGSPGASQIEKLFEDY 343
 DB 365 SNFYAVNLTKEENSKIRYFMIDYDMSAPNVEYEID---ANNKNPYSSEELKELYDY 420
 QY 344 YDSIDLPIHPOFGSRDYEAFLNGIPSGGLFTGAGGIMSEENASRMGGAAGVADANY 403
 DB 421 YKAHHLNTLTPFDGRSDYVGFINGIPAGGIAIGAE---KNVNVN-----GKYLDCRY 471
 QY 404 HAAGDNMTNLNHEAFLINSKATAFNAVATYANDLSSIPKRRNT 444
 DB 472 HQLCDVSNLSMDAFTITKLIASHVATYADSPFEPKRET 512

RESULT 6
 Y04867
 ID Y04867 standard; Protein; 500 AA.
 AC Y04867;
 DT 06-JUL-1999 (first entry)
 DE Mycobacterium species protein sequence 19D.
 KW Secreted protein; Mycobacterium; primer: PCR; amplification; probe:
 KM hybridisation; detection; vaccine; immunisation; infection.
 OS Mycobacterium sp.
 PN M09090186-A2.
 PD 25-FEB-1999.
 PF 14-AUG-1998; F01813.
 PR 11-SEP-1997; FR-011325.
 PR 14-AUG-1997; FR-010404.
 PA (INSP) INST PASTEUR.
 PI Gicquel B, Lim EM, Pellicc V, Portnoi D, Goguet de la Salmoniere Y,
 PI Guilgueno A; 99-181045/15.
 DR N-PSDB; X34119.
 PT Mycobacterial DNA vectors containing reporter constructs - for

PT identifying coding or promoter sequences involved in
 PS infection-associated protein expression
 CC Claim 32; Fig 19D; 305pp; French.
 CC Sequences Y04742-Y05000 and Y07201-Y07204 represent secreted proteins
 CC from various *Mycobacterium* species microorganisms. The encoding
 CC nucleotide sequences can be used as primers and probes for methods
 CC for detecting and identifying mycobacteria, especially belonging to
 CC the *M. tuberculosis* complex. The encoded proteins can be used in
 CC vaccines for immunisation against a bacterial or viral infection.
 CC Sequence 500 AA;

Query Match 26.7%; Score 668; DB 1; Length 500;
 Best Local Similarity 38.5%; Pred. No. 2.6e-50;
 Matches 170; Conservative 57; Mismatches 184; Indels 30; Gaps 11;

QY 14 IQLEDLESQLEDFAYAYPERNRVFGRAHDDYNYLYEELKGTGYDYVYKQPOVHLM 73
 DB 52 VSTDAMAHLSKLQDIDANA-NDGTRAVGTPGYQASDYVYVNTLRNSGFDVQTEFSARVF 110
 QY 74 SMDQTLKVGDEIEATMTYS-----PSYETADYAVYKNLGCSEADYPS-DYE 122
 DB 111 KAEKVYTLGENTYEAHALYSLSGTPPDGYTGPIVAAAPD---DSPGCSDDYRLPVS 166
 QY 123 GKVALIRGCEPFGDKSVLAAKAKAAAIYNNV-AGSMAGTIGAASQDPYSATYGIS 181
 DB 167 GAVYLVDRGCPFGKQKDAAGAVALLIADNIDEGAMGTIGANTDVKIP---YVSVT 223
 QY 182 LEDGQKILKLAEGSVSVLDLVWSKQENRTYVNVVQTKGDPNNVVALGHTDSVEAGP 241
 DB 224 KSVGFQI--RGOSGPTVVKL--TASTQSFARVNIQTKGSSANVYMAHAHDSVEGP 279
 QY 242 GINDGSGIISNLYAKAL-TQYSVKAVERLFMTAEEFGLSNYVSHNATELNKIR 300
 DB 280 GINDSGGVAVLETAVALQGNRPVSNVAFARFAGAEFEGLISRNYESLDLALGIA 339
 QY 301 LYLFNFMASPNYALMTYDGGSAFNOGSG-----PASQAIEKLFEDYDSDILPHIPTO 355
 DB 340 LYLFNFMASPNYVETAYVQGNRPVSNVAFARFAGAEFEGLISRNYESLDLALGIA 399
 QY 356 FDGASDYEAFLNGLIPSGGLFTGAGGIMSEENASRMGGAAGVADANYHAAGDNMTNLN 415
 DB 400 FDGASDYEAFLNGLIPSGGLFTGAGGIMSEENASRMGGAAGVADANYHAAGDNMTNLN 459
 QY 416 EAFLINSKATAFNAVATYANDL 436
 DB 460 TALGINSAGVAYAGLYADL 480

RESULT 7
 Y04868
 ID Y04868 standard; Protein; 527 AA.
 AC Y04868;
 DT 06-JUL-1999 (first entry)
 DE Mycobacterium species protein sequence 19F.
 KW Secreted protein; Mycobacterium; primer: PCR; amplification; probe:
 KM hybridisation; detection; vaccine; immunisation; infection.
 OS Mycobacterium sp.
 PN M09090186-A2.
 PD 25-FEB-1999.
 PF 14-AUG-1998; F01813.
 PR 11-SEP-1997; FR-011325.
 PR 14-AUG-1997; FR-010404.
 PA (INSP) INST PASTEUR.
 PI Gicquel B, Lim EM, Pellicc V, Portnoi D, Goguet de la Salmoniere Y,
 PI Guilgueno A; 99-181045/15.
 DR N-PSDB; X34120.
 PT Mycobacterial DNA vectors containing reporter constructs - for
 PT identifying coding or promoter sequences involved in
 PS infection-associated protein expression
 CC Sequences Y04742-Y05000 and Y07201-Y07204 represent secreted proteins

CC from various Mycobacterium species microorganisms. The encoding
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
SQ Sequence 527 AA;

Query Match 26.7%; Score 668; DB 1; Length 527;
Best Local Similarity 38.5%; Pred. No. 2.8e-50;
Matches 170; Conservative 57; Mismatches 184; Indels 30; Gaps 11;

QY 14 IQLELLSGSQLEDPATAYPERNRVFGKADDTVNYIYEELKTKGYDYV KQPOVHLM 73
DB 79 VSTDMAHMLSLQDIANV-NDGTRAVGTPGYQASVDYVNTLRNSGPEVQYPEFSARVF 137
QY 74 SNADQTLKVGDEIEAKTKTYS-----PSYEVTADAVVKNLGCSEADYPS-DYE 122
DB 138 KAEKGVVILGNTVAKALEYSLGTPPDGVTGPLYAAPAD---DSPGCSPIIDRLPVS 193
QY 123 KVALIKRGCEPFGDKSVLAAKAKAASIVYNNV-AGSNAGTLGAOSDKGYSATVGS 181
DB 194 GAVVLVDRGVCFFAQKEDAAAGVALLIADNIDQAMGTLGANTDKIP--VVSVT 250
QY 182 LEDGKTLKLAAGSVSDLVWDSKQENRTTYNVAQOTGSGDPNNVVALGHTDYEAGP 241
DB 251 KSVGFQI--RGSGPTTYKL--TASTQSKARNVIAQOTGSSAVVYVAGAHLDPEGP 306
QY 242 GINDSGSIISLTVAKAL-TOYSKNAVRFLEMTAEFGILGSNYVYSHLNATELNKR 300
DB 307 GINDSGSVAAALERTAVQVGNPHYSNVRFAFWGAEFGILGSNYVYSHLIDIALKGIA 366
QY 301 LYLNDMTIASPNYALMTYDGGG-AFNOSG---PAGSAQIEKLFEDYDSDIDPHIPTQ 355
DB 367 LYLNDMTIASPNYGFYTDGQSLDPLDARGQVYVEGSGIERTVAYILKAGKTAQDTS 426
QY 356 FDGREDYEAFLINGIPSGLFTGAGIMSEENASRWGQAGAVYDANYHAADNMTNLNH 415
DB 427 FDGREDYGFYTAGIPSGLEFGAIVKKSQAQELMGTADPEPDPNHTQKTDITLHDH 486
QY 416 EAFINSKATAFAVATYANDL 436
DB 487 TALGINGAGVAVAGLYAQL 507

RESULT 8

ID W85456 standard; Protein; 472 AA.
AC W85456.

DE 25-FEB-1999 (first entry)
DB Secreted protein encoded by clone bu45_2.
KW Suppressing activity; nutritional activity; immune stimulating; vaccine;
KW Suppressing activity; haematopoiesis regulating activity;
KW tissue growth activity; activin; inhibin activity; chemotaxis;
KW chemokine activity; haemostasis; thrombolytic activity; receptor;
KW ligand; anti-inflammatory; cadherin; tumour invasion suppressor;
KW tumour inhibition; gene therapy.
OS Homo sapiens.
FN M09842739-A2.
PD 01-OCT-1998.
PF 20-MAR-1998; U05653.
PR 19-MAR-1998; US-044466.
PR 21-MAR-1997; US-822167.
PA (GENY) GENETICS INST INC.
PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
PI Racie LA, Spaulding V, Treacy M;
DR WPI: 98-609890/51.
DR N-PSDB: W82779.
DR New polynucleotides encoding secreted human proteins - derived from
PT human foetal brain, adult brain, foetal kidney, placenta or adult
PT pineal gland cDNA libraries.
PS Claim 14: Page 70-72; 113pp; English.
CC The present sequence represents a secreted protein. The polynucleotide

CC and secreted protein are predicted to have biological activities which
CC would make them suitable for treating, preventing or ameliorating medical
CC conditions in humans and animals, although no supporting data is given.
CC Suggested activities include nutritional activity, immune stimulating
CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokine activity, haemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
CC invasion suppressor activity, and tumour inhibition activity (no data is
CC given in the specification to support these activities). The
CC polynucleotide is also stated to be useful for gene therapy.
SQ Sequence 472 AA;

Query Match 9.2%; Score 229.5; DB 1; Length 472;
Best Local Similarity 22.6%; Pred. No. 5.4e-12;
Matches 96; Conservative 22, Mismatches 175; Indels 71; Gaps 19;

QY 38 RVFGKADDTVNYIYEELKTKGYDYV KQPOVHLSNADQTLKVGDEIEF-----AK 90
DB 77 RLSGSKNKEKAIQIYQVNLQODGLEKVLLEPRIPMERGESAVMLPRIKHIALIG 136
QY 91 TMTYPSYEVTADAVVKNLGCSEADYPSDYGKVALIKRGCEPFGDKS-----Y 140
DB 137 SGTGTPGEGITREVLVVSFDELQ-RASEARKIIVINQ---PIINSTRYQYRTQAV 192
QY 141 LAAKAKAASIVYNNVAGSMAGTLGAO--SDKGPYSATVGSLEDGKTLKLAAG-SV 197
DB 193 EAKAGALASILRSVASISVSPHNGIQEDYDGVKIPACTIVADAMSMASHGKI 252
QY 198 SVDLWDSKQENRT-TYNNVAQOTKGD-PNNVVALGHTDYEAGP GINDSGSIISLV 255
DB 253 VIQLMGAKTYPPDDISFTVAETIGSKYPEQVYVLSGHLDSMDVGQGMDDGGGFIISWE 312
QY 256 IAKALTQSV--KNAVRFLEMTAEFGILGS-NYVSHLNATELNKRILYINFDIASPN 312
DB 313 ALSLKDKGLPKRKLRLVMTAEBOGVGAFQY-----QLHKVNI-----SN 356
QY 313 YALMTYDGGSAFNOSGP--AGSAQIEKLFED--YDSDIDPHIPTQFDRSDYEATIL 367
DB 357 YSL-VNESDAGFLPTGQFTGSEKARAIMEEVMSLLQPLNTQVLSHGEG-TDINFWIQ 414
QY 368 NGIPSGGLFTGAGIMSEENASRWGQAGVAYDANYHAADNMTNLNEAFINSKATAF 427
DB 415 AGVPASLIDL-----YKFFHHSHGDTMTVMDPKQ--MNVAAVW 455
QY 428 AVATY 432
DB 456 AVVST 460

RESULT 9

ID W33604 standard; Protein; 472 AA.
AC W33604.

DE 21-MAY-1998 (first entry)
DB Human secreted protein AM282 full-length sequence.
KW Secreted protein; AM282; cytokine; human.
OS Homo sapiens.
FN Key Location/Qualifiers
FT Peptide 1..24
FT /label= Sig_peptide
FT /label= 25..472
FT Protein /label= Mat_protein
FN W09739030-A2.
PD 23-OCT-1997.
PF 16-APR-1997; U06475.
PR 13-JAN-1997; US-783520.
PR 18-APR-1996; US-634325.
PA (GENY) GENETICS INST INC.
PI Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA,
PI Spaulding V;
DR WPI: 97-526400/48.

DR N-PSDB:V02296.
PT New isolated secretory proteins AM340, AM282 and AK583 - possibly
PT have cytokine, cell proliferation/differentiation regulating,
PT immunomodulating activities, etc.
PS Claim 16; Page 45-47; 59pp; English.
CC This human secreted protein, designated AM282, is encoded by a
CC full-length cDNA clone (see V02296), deposited in ATCC 98026, that
CC was identified from a database search using an isolated partial
CC AM282 clone (see T97398). AM282 protein can be used in a claimed
CC method for preventing, treating or ameliorating a medical
CC condition. It may exhibit cytokine, cell proliferation (either
CC inducing or inhibiting) or cell differentiation (either inducing or
CC inhibiting) activity or may induce production of other cytokines in
CC certain cell populations. It may also exhibit e.g. immune
CC stimulating or suppressing activity, haematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic or chemokinetic activity, haemostatic or thrombolytic
CC activity, receptor/ligand activity, anti-inflammatory activity,
CC tumour inhibition activity, or other activities. No evidence of
CC any of these activities is given in the specification.
SO Sequence 472 Aa:

[illegible]

PR 19-MAR-1991: US-671376.
PA (OMNT-) OMNIGENE, INC..
FI Pero J, Ruto GA, Sloma A;
DR WPI: 92-349223/42.
N-PSDB: Q29134.
PT Bacillus residual protease III (Rp-III) gene - which can be
PT mutated for prodn. of heterologous polypeptide(s) or expressed to
PT produce protease
PS Claim 25: Fig 4: 28pp: English.
CC This sequence represents residual protease III (rp-III) from
CC Bacillus subtilis, and was decoded from the appropriate DNA as detailed
CC in Q29134. The protease is secreted by B. subtilis in large amounts, and
CC can cause proteolytic degradation of any protein produced by the
CC bacteria, i.e. a recombinant product. By introducing mutations in the
CC protease gene heterologous polypeptides can be expressed which do not
CC cause lowered levels of proteolytic damage. Also Rp-III protease can be
CC expressed for use in improving the cleaning activity of laundry powders
CC or for use in industrial processes.
SQ Sequence 806 AA;

	Query Match:	7.5%;	Score 188.5;	DB 1,	Length 806;
	Best Local Similarity	23.0%;	Pred. No. 4.7e-08;		
	Matches 94;	Conservative 53;	Mismatches 154;	Indels 107;	Gaps 16;
QY	AKTTTSPSYSEVA-----DVAVKMLGCGSEA-DYP-SPECKVALLIRKECFEDSKSVLA	142			
Db	AKWGYRKEDDVKALNNKREVELVE-AGIGEKAKDFEGKNLTGRVAVVRKGSIAPFADKADNA	421			
QY	143 AKAKAAASIVYNNAVSGSMAGTLGAQAOSDKGPYSALVGISLEDGKKLIKLAEGSVS----	199			
Db	422 KKAGAIIMGVNNTLSGEIEANV-----PGMSVPTRIKLSDGEEKLVSAALKAGEFTKTTFK	475			
QY	199 -----VDLMWDKSQENRTTYNYVAQTKGGDPNNNVVALGCHDSYE	238			
Db	476 LTVSKALGCEVADFSSRRGPMDTMIRKPDISAPGVNIYSTIPTHDDPHYGYSKOGTSN	535			
QY	239 AGPINDDGSGIISNVIKALTOXSVKNVAFLEFWTAEEFGLLGSNXYVSHLNATELNK	298			
Db	536 ASPIH-----AGAVA--VIQAKPKPSVEDI-----	560			
QY	299 IRLYNFDMIASPVALMITYDDGSAF---NOSGPAGSAQIEKLEDEYYDSIDLPHIPTOP	356			
Db	560 -----KAIMINTRAVTCLKDSGDGEVYFHNNQG-AGSARINNALK-ADSLVSP-----	603			
QY	357 DGRNDYEAFLNGICIPSGGLFTGAEIGINSEENNSRKGQAGVAYANTHAANGDMTNLNEH	416			
Db	603 -GSYSYGF----LKENGNETKNEFTFIENOSS-----IRKSYLLEYFSFNKSGISTSGTS	652			
QY	417 AFLINSKATFAVAITANDLSSIPRNRTTSSLHRRARTMRFCGRRAK	464			
Db	653 RVVIPAHQTGKATAKY-----KVNRKTKACGTGEYIVRGSGTVAK	695			
RESULT	11				
ID	W31524 standard; protein: 751 AA.				
AC	W31524;				
DT	14-MAY-1998 (first entry)				
DE	Prostate-specific membrane antigen.				
KW	Prostate-specific membrane antigen; PSMA; PSM'; monoclonal antibody;				
KW	prostatic cancer; metastasis.				
OS	Homo sapiens.				
FH	Key				
FT	Protein				
FT	Location/Qualifiers				
FT	/label= PSM'				
FT	/note= "Putative PSM' protein, a variant of PSMA"				
FT	Domain				
FT	/note= "Putative transmembrane domain"				
FT	Region				
FT	/note= "peptide selected for monoclonal antibody development"				
FN	W09735616-A1.				

PD 02-OCT-1997; U05214.
PF 25-MAR-1997; U05214.
PR 25-MAR-1996; U05214399.
PA (PACI-) PACIFIC NORTHWEST CANCER FOUND.
PI Boynton AL, Holmes EH, Murphy GP, Tino WT;
DR WPI: 97-489396/45.
PT Monoclonal antibody against prostate-specific membrane antigen -
PS used for diagnosis and treatment of prostatic cancer
PS Claim 2; Fig 1; 76pp; English.
CC The present sequence represents the prostate-specific membrane antigen
CC (PSMA). A novel monoclonal antibody (MAb) has been developed which has
CC an antigen-binding region specific for the extracellular domain (ECD) of
CC PSMA. The MAb is used for detecting the presence of PSMA (or its new
CC variant PSM', which lacks the intracellular domain (ICD)) in biological
CC specimens or cancer cells, particularly for the diagnosis, prognosis and
CC monitoring of prostate cancer, in which levels of PSMA are elevated.
CC Kits for carrying out this method are also included within the scope of
CC the invention. The MAb can also be used for treatment of prostatic
CC cancer and associated metastases, optionally when conjugated to a drug,
CC toxin or radioisotope. MAb directed against ECD provide a more
CC effective detection than the known antibody 7E11-C15 which binds to the
CC ICD, and detects PSMA only in necrotic or apoptotic cells (and can not
CC detect PSM' at all). Tests for PSMA can be done on blood and urine,
CC without the need for a biopsy sample.
SQ Sequence 751 AA;

Query Match 6.4%; Score 161; DB 1; Length 751;

Best Local Similarity 19.8%; Pred. No. 1.1e-05;
Matches 96; Conservative 64; Mismatches 183; Indels 142; Gaps 17;

QY 18 DLLESGQLEDAFAYAPENRRPFGCAHDDYVNYLYEELKKTG-----YDYV----- 66
DB 63 DLELAENKTKFLYNTQIPHLAGEQNFOLAKIQSOQKKEGLDSVEIAHYLLSTYPN 122
QY 66 -KOPVHLSNADQTLKVE-----EIEAKTWTSPSEVETADVAV 106
DB 123 KHPNVTISINED-----GNEIFNTSLFEPFPGYENVDIYPPFAFPGGPEDDLYV 177
QY 107 V-----KNIGCEADYPSVDEKVALIKRGCEPFQDKSVLAKAKAAHIVYNNA----- 158
DB 178 VYVARTEDFEKLERDKIKNSGKIVTARYKVGKKNQAQAGKGVILXSDPADYFA 237
QY 158 -----GSMAGTL-----GAASDCKGPIYAIVG 179
DB 238 PGVKSYPDGMNLPGGGVORGNILNNGAGDPLTPGYRANEYVRRGIAVGLPSIPVAP 297
QY 180 ISLEDOQKLIK-----LAAGSVSYDLWVDSKQENRTTY 213
DB 298 IGYDAQKLEKMGAGAPDSSMRGSLKYPYVNGPGFTGNFSTQKYMHIHSTNEVTRIV 357
QY 214 NVVAOTKGG-DENNVVALGHTDSYVAPGIN-DDSGSIISMLVIAKALTQ---YSVKNA 268
DB 358 NVIGTLRGAVEDRYILGHRDSWYFG-GIDPQGAAYVHEIVRFGLTKREGWPRRT 416
QY 268 VEFLEWTAEEFLLGNSNYVSHLNATELKKIRLYLNFDMIASPNYALITYDDGSGAFNOS 328
DB 417 ILFASMDAEEFLLGSTEVAEENSRLLOERGAVINADSSIGENWYLFV---DCPPLMYS 473
QY 329 -----GPAGSAQIEKLFEDYDYSIDLPHIPTOFDGR-----SDYEAFILN- 369
DB 474 LVHNLTKELKSPDEGEKSLYESWTKKSPSP-----EFGSMRISIKLSGNDPEVYFQAL 529
QY 369 GIPSG 373
DB 530 GIASG 534
RESULT 12
R99416
ID R99416 standard; Protein; 393 AA.
AC R99416;
DT 04-DEC-1996 (first entry)

DE Amino-peptidase precursor of Aeromonas caviae, useful in food mfr.
KW Amino-peptidase; hydrophobic amino acid removal; hydrolysis;
KW food manufacture; improved taste.
OS Aeromonas caviae T-64.
FH Key Location/Qualifiers
FT protein 121..393
FT /label=mat_peptide
PD J08173168-A.
PD 09-JUL-1996.
PF 26-DEC-1994; 336663.
PR 26-DEC-1994; JP-336663.
PA (NORO) NORINSUISANSO SHOKUHN SOGO.
PA (ZENK-) ZENKOKU RAKUNOGYO KYODO KUMIAI RENOGOKAI.
DR WPI: 96-365587/37.
DR N-PSDB: T41568.
PT DNA encoding aeromonas microorganism derived amino-peptidase - which
PT eliminates bitter tastes by decomposing peptide(s) rich in
PT hydrophobic amino acids produced by protein hydrolysis, useful in
PT food mfr.
PS Claim 1; Page 5-7; 12pp; Japanese.
CC R99416 is an amino peptidase precursor from Aeromonas caviae. The
CC peptidase is useful in the hydrolysis of peptides high in hydrophobic
CC amino acids, this is useful in food manufacture as such peptides have
CC a bitter taste. The amino-peptidase removes amino acids sequentially
CC from the N-terminus of a peptide at an optimum reaction temperature
CC of 50 deg. C. and a pH of between 8 and 10.
SQ Sequence 393 AA;

Query Match 6.3%; Score 157.5; DB 1; Length 393;

Best Local Similarity 25.2%; Pred. No. 8.2e-06;
Matches 78; Conservative 37; Mismatches 113; Indels 81; Gaps 17;

QY 158 GSMAGTLGAQSDKGP-YSAIVGISLED---GQKLKIAAGSVSYDL-WVDSKQENRTT 212
DB 123 GNIVGTISQLASMRNRYTTTGGVQADWVAGQ-----WQSLATLPMWASVRYKKG 175
QY 213 Y---NVVAOTKGG-DENNVVALGHTDSYEA-----GPEINDGSIISMLVIAKAL 260
DB 176 YPQGSVVLTLKSSRYPDEVVYVGLGHDSTPAGSAPNSRLAPADDAAGIATLTVLRVI 235
QY 261 TGY---SVKNAVRELPWTAEEFLLGNSNYVSHLNATELKKIRLYLNFDMIASPNYALMIX 318
DB 236 AEOGRPEKTLPIGYAAAEVGLRSGSKDIATRYKAAK-TKVLAALQIDM---TNY----- 287
QY 319 DGDGSAFNOSGPAGSAQIEKLFEDYDS-----ID-LPHIPTOFD---GRSDYE 363
DB 287 -----GQSAEDIVFMTDYDTDQGLTGYLAQLLDAYLPQIRYGVDSGCGSDHA 334
QY 364 AFLINGIPSGGLFTGAEGLMSEENASRMGCGAGVADANYHAAGDMNTLN---HEAFIL 420
DB 335 SWHNOGYPRAMPF-----ESRFND-----YNPRIHTAODTLONSDPSAAHALKF 378
QY 421 NKKATAFAV 429
DB 379 AQLATSFAT 387
RESULT 13
W47155
ID W47155 standard; Protein; 707 AA.
AC W47155;
DT 28-MAY-1998 (first entry)
DE Extracellular domain of prostate specific membrane antigen (PSMA).
KW Fc region; immunoglobulin; recombinant DNA; target protein; PSMA;
KW fusion protein; prostate specific membrane antigen; immunofusion;
KW extracellular domain.
OS Homo sapiens.
PN US5726044-A.
PD 10-MAR-1998.
PF 14-SEP-1995; 528122.
PR 14-SEP-1995; US-528122.
PR 14-SEP-1994; US-305700.

PA (FUT-1) FUJI IMMUNOPHARMACEUTICALS CORP.
 PI Gallies SD, Lo K, Sudo Y;
 DR WPI: 98-192768/17.

PT Recombinant DNA for expression of target protein, e.g. HIV gp120 -
 PT comprises sequences coding for signal peptide, immuno-globulin Fc
 region and gp120

PS Example 14: Columns 27-32: 18pp; English.

CC This is the extracellular domain of the prostate specific membrane
 CC antigen (PSMA) which is a type II membrane protein. The extracellular
 CC domain is modified to construct a new recombinant fusion protein, a
 CC PSMA immunofusion. The new recombinant DNA construct is used for
 CC expression and secretion of a target protein, whose sequence is free of
 CC immunoglobulin CH1 domain. The construct comprises a polynucleotide
 CC encoding from 5' to 3' direction a secretion signal sequence,
 CC comprising a sequence encoding an immunoglobulin Fc region, and a
 CC sequence encoding the target protein. A replicable expression vector
 CC comprising the polynucleotide sequence can be used to transfect a host
 CC cell. The products can be used to produce a recombinant fusion protein
 CC (immunofusion) comprising the Fc region and target protein. The DNA can
 CC be expressed at high levels in a host cell, and the fusion protein is
 CC efficiently produced and secreted.

SQ Sequence 707 AA;

Query Match 6.28; Score 154; DB 1; Length 707;

Best Local Similarity 20.18; Pred. No. 4e-05;
 Matches 90; Conservative 59; Mismatches 170; Indels 128; Gaps 17;

QY 28 DFAAYAPERNRVFGKADHDVNYLYE--ELKRTGYDYVKOPQVHLSNADQTLKVG 84
 DB 71 DVLSTYF-----NKRHPYISINEDGNEIFNTSLF-----EPPPGYENV----- 112
 QY 85 EEIFAKTMTYSPSYEVYADAVV-----KNLGCSEADYPSDVEGKVALIKRGCEPFGDKS 139
 DB 112 SDIVPPSAFSPQMGEGDLVYVYARTEDFELERDKKINCCKGKVIARYGKVFGRNKV 171
 QY 140 VLAAKAKAAAIYVNNVA-----GSMAGTL----- 165
 DB 172 KNAQLAGAKGYILSDPADYFAPGVKSYPRDGMNLPGGGVQNGNITLNGAGDPLTPGYRA 231
 QY 165 -----GAOSDKGPRYSATVIGISLEDGQKLIK-----L 191
 DB 232 NEYAYRRGIAEAVGLPSIPVHPIGYDAQKLEKMGKSAPPDSWRGSLKVPYVNGPGFT 291
 QY 192 AEAAGSVVDLWDSKQENRTYVNVVQTKGG-DPNNVVALGHTDSVAGGIN-DDGSG 249
 DB 292 GNESTQVKKMHISTNEVTRITVYIGLRGAVEPRDRIYILGHRDSWVFG-GIDPOSGAA 350
 QY 250 IISNLVIKALTO--YSVNAVRFLFWTAEEFGILGSNYYVSHLNATELNKIRLYLNF 306
 DB 351 VVHEIYVNSFGTLKEGMRPRTILFASMDAEFGILGSTMAEENSRLOERGVAYINAD 410
 QY 307 MIASPNYALMIYDGDGSAFNOS-----GPAGSAQIEKLFEDYDSDIDPHIPIQF 356
 DB 411 SSIEGNTYLRV---DCRPLMYSLVHNLTRELKSPDEGFEKSLYESWTKKSPSP---EF 463
 QY 357 DGR-----SDYEAFILN-GIPSG 373
 DB 464 SGMPRIKSLGSGNDFEVEFQRLGIASG 490

RESULT 14
 R55097
 ID R55097 standard; Protein; 750 AA.
 AC R55097;
 DT 11-JAN-1995 (first entry)
 DE Prostate-specific membrane antigen.
 KW Prostate-specific membrane antigen; PSM; prostate cancer;
 KW transmembrane glycoprotein; imaging; targeting; tumour detection;
 KW antibody detection.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 63.68

FT peptide /note- "region of high hydrophilicity 1, claim 35"
 FT peptide 132..137
 FT peptide /note- "region of high hydrophilicity 2, claim 35"
 FT peptide 482..487
 FT domain /note- "region of high hydrophilicity 3, claim 35"
 FT domain 23..39
 FT domain /note- "specific membrane spanning domain"

PN W09409820-A.
 PD 11-MAR-1994.
 PE 05-NOV-1993; U10624.
 PR 05-NOV-1992; US-973337.
 PA (SLOK) SLOAN KETTERING INST CANCER.
 PI Falt WR, Heston MDW, Israel RS;
 DR WPI: 94-167129/20.

DR N-PSDB; 065520.
 PT Prostate-specific membrane antigen and DNA encoding it - is
 PT useful for detecting haematogenous micro-metastatic tumour cells
 PT and for identifying ligands which bind to PSM Ag

PS Claim 24: Page 103-106; 196pp; English
 CC R55097 is a prostate specific membrane antigen (PSM, encoded by
 CC 055520). The inventors have identified 3 peptides of high
 CC hydrophilicity to which antibodies may be directed against. In the
 CC specification the specific membrane spanning domain is stated to be
 CC from amino acid residue 19 to residue 44. The PSM coding sequence
 CC is useful for suppressing or modulating the metastatic ability of
 CC prostate tumour cells to grow, or for eliminating them. The protein
 CC is useful to identify or purify ligands of the Ag. It is also an
 CC attractive target for Ab-directed imaging and targeting of prostatic
 CC tumour deposits.

SQ Sequence 750 AA;

Query Match 6.28; Score 154; DB 1; Length 750;

Best Local Similarity 20.18; Pred. No. 4.4e-05;
 Matches 90; Conservative 59; Mismatches 170; Indels 128; Gaps 17;

QY 28 DFAAYAPERNRVFGKADHDVNYLYE--ELKRTGYDYVKOPQVHLSNADQTLKVG 84
 DB 114 DVLSTYF-----NKRHPYISINEDGNEIFNTSLF-----EPPPGYENV----- 155
 QY 85 EEIFAKTMTYSPSYEVYADAVV-----KNLGCSEADYPSDVEGKVALIKRGCEPFGDKS 139
 DB 155 SDIVPPSAFSPQMGEGDLVYVYARTEDFELERDKKINCCKGKVIARYGKVFGRNKV 214
 QY 140 VLAAKAKAAAIYVNNVA-----GSMAGTL----- 165
 DB 215 KNAQLAGAKGYILSDPADYFAPGVKSYPRDGMNLPGGGVQNGNITLNGAGDPLTPGYRA 274
 QY 165 -----GAOSDKGPRYSATVIGISLEDGQKLIK-----L 191
 DB 275 NEYAYRRGIAEAVGLPSIPVHPIGYDAQKLEKMGKSAPPDSWRGSLKVPYVNGPGFT 334
 QY 192 AEAAGSVVDLWDSKQENRTYVNVVQTKGG-DPNNVVALGHTDSVAGGIN-DDGSG 249
 DB 335 GNESTQVKKMHISTNEVTRITVYIGLRGAVEPRDRIYILGHRDSWVFG-GIDPOSGAA 393
 QY 250 IISNLVIKALTO--YSVNAVRFLFWTAEEFGILGSNYYVSHLNATELNKIRLYLNF 306
 DB 394 VVHEIYVNSFGTLKEGMRPRTILFASMDAEFGILGSTMAEENSRLOERGVAYINAD 453
 QY 307 MIASPNYALMIYDGDGSAFNOS-----GPAGSAQIEKLFEDYDSDIDPHIPIQF 356
 DB 454 SSIEGNTYLRV---DCRPLMYSLVHNLTRELKSPDEGFEKSLYESWTKKSPSP---EF 506
 QY 357 DGR-----SDYEAFILN-GIPSG 373
 DB 507 SGMPRIKSLGSGNDFEVEFQRLGIASG 533

RESULT 15
 W02234
 ID W02234 standard; Protein; 750 AA.
 AC W02234;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2000, 19:06:39 ; Search time 31.24 seconds
(without alignments)
222.696 Million cell updates/sec

Title: US-09-080-127-2_COPY_16_496

Perfect score: 2500

Sequence: 1 GRALVSPDFEPEDIQLEDLI.....APKTHAHVSGSGCWHSQVERA 481

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 143561 seqs, 14463640 residues

Total number of hits satisfying chosen parameters: 143561

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued Patents-AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/PTCUS.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	154	6.2	707 1	US-08-528-122-18
2	154	6.2	707 4	PCT-US95-11720-18
3	154	6.2	750 1	US-08-325-553-2
4	154	6.2	750 2	US-08-394-152A-2
5	149	6.0	1167 2	US-08-589-756-2
6	148	5.9	1150 2	US-08-589-756-3
7	145	5.8	1164 2	US-08-589-756-1
8	140.5	5.6	622 1	US-08-547-197-1
9	134.5	5.4	372 1	US-08-929-922B-2
10	130	5.2	760 1	US-08-547-197-2
11	109	4.4	1026 2	US-08-542-003-6
12	109	4.4	1026 2	US-08-323-760A-6
13	105.5	4.2	551 2	US-09-033-537A-1
14	103	4.1	640 2	US-08-671-978A-10
15	102.5	4.1	1545 3	PCT-US95-10661A-4
16	101	4.0	1049 3	US-08-772-270A-11
17	101	4.0	1244 4	PCT-US93-10500-2
18	100	4.0	628 1	US-08-261-832A-6
19	100	4.0	628 4	PCT-US95-0744A-6
20	99	4.0	687 1	US-08-204-656B-10
21	99	4.0	687 1	US-08-470-702-10
22	99	4.0	687 1	US-08-467-831-10
23	99	4.0	928 1	US-08-474-140-11
24	99	4.0	928 1	US-08-477-630-11
25	99	4.0	928 1	US-08-477-293-11
26	99	4.0	928 1	US-08-474-345-11
27	99	4.0	928 1	US-08-478-341-11
28	98.5	3.9	1794 5	5183745-6
29	97.5	3.9	503 1	US-07-612-673-2

30	97.5	3.9	521 1	US-07-796-361A-11	Sequence 11, Appl
31	97.5	3.9	521 1	US-08-539-666-2	Sequence 2, Appl1
32	96	3.8	566 1	US-08-415-823-4	Sequence 4, Appl1
33	96	3.8	566 2	US-09-086-662-4	Sequence 4, Appl1
34	96	3.8	715 2	US-08-849-212-6	Sequence 6, Appl1
35	95.5	3.8	1222 2	US-08-682-517-15	Sequence 15, Appl
36	95.5	3.8	1252 2	US-08-682-517-9	Sequence 9, Appl1
37	95.5	3.8	4350 2	US-08-804-227C-8	Sequence 8, Appl1
38	95.5	3.8	4550 2	US-08-804-198-2	Sequence 2, Appl1
39	95	3.8	1529 2	US-08-728-470-10	Sequence 10, Appl
40	95	3.8	1600 2	US-08-617-697-10	Sequence 4, Appl1
41	93	3.7	318 2	US-09-027-013-4	Sequence 2, Appl1
42	93	3.7	820 1	US-08-261-896-2	Sequence 2, Appl1
43	93	3.7	820 2	US-08-485-278-2	Sequence 2, Appl1
44	93	3.7	997 2	US-08-387-942C-4	Sequence 4, Appl1
45	93	3.7	1157 1	US-07-876-280-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-08-528-122-18
: Sequence 18, Application US/08528122
: Patent No. 5726044
: GENERAL INFORMATION:
: APPLICANT: LO, KIN-MING
: APPLICANT: SUDO, YUKIO
: APPLICANT: GILLES, STEPHEN D.
: TITLE OF INVENTION: EXPRESSION AND EXPORT TECHNOLOGY OF
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
: ADDRESSEE: THIBEAULT
: STREET: 125 HIGH STREET
: CITY: BOSTON
: STATE: MA
: COUNTRY: USA
: ZIP: 02110
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/528,122
: FILING DATE:
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: PITCHER, EDMUND R.
: REGISTRATION NUMBER: 27,829
: REFERENCE/DOCKET NUMBER: FIP-001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-248-7000
: TELEFAX: 617-248-7100
: INFORMATION FOR SEQ ID NO: 18:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 707 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: FEATURE:
: NAME/KEY: Protein
: LOCATION: 1..707
: OTHER INFORMATION: /note="EXTRACELLULAR DOMAIN OF
: OTHER INFORMATION: PSMA"
: US-08-528-122-18
: Query Match 6.2%; Score 154; DB 1; Length 707;
: Best Local Similarity 20.1%; Pred. No. 2.2e-06;

APPLICATION NUMBER: US/07/973.337A
FILING DATE: 05 NOV 1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 1747/41426
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 750 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-325-553-2

Query Match 6.2%; Score 154; DB 1; Length 750;

Best Local Similarity 20.1%; Pred. No. 2.5e-06;
Matches 90; Conservative 59; Mismatches 170; Indels 128; Gaps 17;

QY 28 DEAYAYPERNNVFGKAKHDTVNTLYE--ELKRTGYDYVKQOVHLSNADQTLKVG 84
DB 114 DVLSTY-----NKHHPNTISINEDGNEIFNTSLF---EPPPGYENV----- 155
QY 85 EEIRAKTMTYSPSYEVTAADVAVV-----KNIGCEADYPSDVEGKVALIRGCEPFGDKS 139
DB 155 SDIYPPSAPSPQGMPEGLDLYVNVARTEDPFKLERDMKINCISKIYIARYKVFGRNKV 214
QY 140 VLAAKAKAASIVYNNVA-----GSMAGTL----- 165
DB 215 KNAOLAGAKGYILYSDPADYFAPGVKSPDGWNLPGGGVQVQGNILNLNGAGDPLTPGYPA 274
QY 165 -----GAOSDKGPYSIAVIGISLEDOGLIK-----L 191
DB 275 NEYAYRNGIAEAVGLPSIPHPITGYDAQKLEKMGSGAPPDSSWRGSLKYPYVVGFT 334
QY 192 AEAGSVSDLVWDKQENRTYNNVAQTKG-DPNNVVALGHTDSVEAGPGIN-DDGSG 249
DB 335 GNESTOKVKMHIHSTNEVTRIYNVIGTLRGAVEDRDYVILGHRDSWVFG-GIDPQSGAA 393
QY 250 IISNLVIAKALQO---YSVKNAVAFLEWTAEEFGILGSNVYVSHLNATELKIILYINFD 306
DB 394 VVHEIVRSFGTLKKEGWRPRTILFASWDAEFGILGSTENAENSRLLOERGVAIYNAD 453
QY 307 MIASPNALMIYDGDGSAFNOS-----GPAGSAQIEKLFEDYDSDIDLPHIPTOF 356
DB 454 SSIGNTTLRY---DCTPLMYSLVHNLTKELKSPDEGEKSLYESWTKKSPSP-----EF 506
QY 357 DGR-----SDYEAFILN-GIPSG 373
DB 507 SGMPRISKLSGNDFEVFFORLGIASG 533

RESULT 4
US-08-394-152A-2
Sequence 2, Application US/08394152A
Patent No. 5935818
GENERAL INFORMATION:
APPLICANT: Israel, Ron S.
APPLICANT: Heston, Warren D.W.
APPLICANT: Fair, William R.
TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America

ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM 330 466 DX2
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,152A
FILING DATE: 24-FEB-95
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41426-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 391-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 750 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-394-152A-2

Query Match 6.2%; Score 154; DB 2; Length 750;

Best Local Similarity 20.1%; Pred. No. 2.5e-06;
Matches 90; Conservative 59; Mismatches 170; Indels 128; Gaps 17;

QY 28 DEAYAYPERNNVFGKAKHDTVNTLYE--ELKRTGYDYVKQOVHLSNADQTLKVG 84
DB 114 DVLSTY-----NKHHPNTISINEDGNEIFNTSLF---EPPPGYENV----- 155
QY 85 EEIRAKTMTYSPSYEVTAADVAVV-----KNIGCEADYPSDVEGKVALIRGCEPFGDKS 139
DB 155 SDIYPPSAPSPQGMPEGLDLYVNVARTEDPFKLERDMKINCISKIYIARYKVFGRNKV 214
QY 140 VLAAKAKAASIVYNNVA-----GSMAGTL----- 165
DB 215 KNAOLAGAKGYILYSDPADYFAPGVKSPDGWNLPGGGVQVQGNILNLNGAGDPLTPGYPA 274
QY 165 -----GAOSDKGPYSIAVIGISLEDOGLIK-----L 191
DB 275 NEYAYRNGIAEAVGLPSIPHPITGYDAQKLEKMGSGAPPDSSWRGSLKYPYVVGFT 334
QY 192 AEAGSVSDLVWDKQENRTYNNVAQTKG-DPNNVVALGHTDSVEAGPGIN-DDGSG 249
DB 335 GNESTOKVKMHIHSTNEVTRIYNVIGTLRGAVEDRDYVILGHRDSWVFG-GIDPQSGAA 393
QY 250 IISNLVIAKALQO---YSVKNAVAFLEWTAEEFGILGSNVYVSHLNATELKIILYINFD 306
DB 394 VVHEIVRSFGTLKKEGWRPRTILFASWDAEFGILGSTENAENSRLLOERGVAIYNAD 453
QY 307 MIASPNALMIYDGDGSAFNOS-----GPAGSAQIEKLFEDYDSDIDLPHIPTOF 356
DB 454 SSIGNTTLRY---DCTPLMYSLVHNLTKELKSPDEGEKSLYESWTKKSPSP-----EF 506
QY 357 DGR-----SDYEAFILN-GIPSG 373
DB 507 SGMPRISKLSGNDFEVFFORLGIASG 533

RESULT 5
US-08-589-756-2
Sequence 2, Application US/08589756
Patent No. 5846547
GENERAL INFORMATION:
APPLICANT: STREPTOCOCCAL C5A PEPTIDASE VACCINE
TITLE OF INVENTION: STREPTOCOCCAL C5A PEPTIDASE VACCINE
NUMBER OF SEQUENCES: 5
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/589,756
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1167 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-589-756-2
```

Query Match 6.0%; Score 149; DB 2; Length 1167;

Best Local Similarity 21.9%; Pred. No. 1.6e-05;

Matches 121; Conservative 66; Mismatches 196; Indels 170; Gaps 26;

```
QY 6 SPDEFEDIQLEDLLEGSQ-----QLEDFAYAPER-----NRFEGK 43
DB 205 ASSETKEPYRLEGAMPEADLLMRVEIYNGLADYARNYAQAIRDAVNLGAKVINMSFGNA 264
QY 44 AHDDTVNT--LYEELKKTGYDYVKOPVYHLSNADOTLK-----YGD 84
DB 265 A---LAYANLPDETKKA--FDYAKSKGVSYTSAGNDSFGKTRPLADHPDYGVGT 318
QY 85 EEIEAKTMT--YSPSVEVTADVAVK-----NLGCS 113
DB 319 PAAADSTILVASISPDKQLT-ETAAVKTTDQODKEMPVLSTNRFEPNKAIDYAIYANRGTK 377
QY 114 EADPSDEGKVALIRGCEPFGDKSVLAAKAAASIYNNVAGSMAGTIGAOSDKG- 173
DB 378 EDFE-KDVAGKIALIRGIDFKDKVAAKAKGAVGLTYDN-----QDKGFP 423
QY 173 -----PYSATYGISLEBQOKLIKLAEGSVSVLDVMSKQENRTTYN--VAOTKGG 222
DB 424 PIELPNVDQMPAAFISSKDG-----LLKDNPKQITTFNATPKVLPFTAG 468
QY 223 DPNNVVALGHDPSVAGGINDGSGIISNLVIK--ALTOYSVKNANRFLFTAEFEG 280
DB 469 KLSRFSWGLTADGNIKRDIAPAGODIISVANNKYAKLSGTSKAPL-----VAGING 523
QY 281 LGSNYYVSHLNTLNLRIYLNFDMIASPNYALMITYDGDSAF--NOSGPAGSAQIEK 337
DB 524 LQKQETQYPMTPSERDLAKKVL-----SSATLALYDEDEKAYSPRQOG-AGAVDAK 578
QY 338 KLF-----EDYDSDIDLPRIPTQFD-----GRSD-----YEAFLNGISPGGLF 376
DB 579 KASAAATMYTDKDNSTSKYHLNNVSDKFEVTVVHNKSDKPOELYQAATVQTDKVDGKHF 638
QY 377 TGAEGIMSE-----ENASRWGQAGVAYDANYHAAGDNMTNLNHEAFILNSKATFA 427
DB 639 ALAPKVLYESMOKITIPANSSK---QYVTPIDAS-RFSKDLAOKKNQGYFL--EGFVAF 692
QY 428 AVATYANDLSIP 440
DB 693 KODPTKEELMSIP 705
```

RESULT 6

US-08-589-756-3

Sequence 3, Application US/08589756

Patent No. 5846547

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: STREPTOCOCCAL C5a PEPTIDASE VACCINE

NUMBER OF SEQUENCES: 5

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

```
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/589,756
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1150 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-589-756-3
```

Query Match 5.9%; Score 148; DB 2; Length 1150;

Best Local Similarity 21.7%; Pred. No. 2e-05;

Matches 120; Conservative 64; Mismatches 200; Indels 168; Gaps 25;

```
QY 6 SPDEFEDIQLEDLLEGSQ-----QLEDFAYAPER-----NRFEGK 43
DB 205 ASSETKEPYRLEGAMPEADLLMRVEIYNGLADYARNYAQAIRDAVNLGAKVINMSFGNA 264
QY 44 AHDDTVNT--LYEELKKTGYDYVKOPVYHLSNADOTLK-----YGD 84
DB 265 A---LAYANLPDETKKA--FDYAKSKGVSYTSAGNDSFGKTRPLADHPDYGVGT 318
QY 85 EEIEAKTMT--YSPSVEVTADVAVK-----YKNLGCS 114
DB 319 PAAADSTILVASISPDKQLT-ETAAVKTTDQODKEMPVLSTNRFEPNKAIDYAIYANRGTK 378
QY 115 ADPSDEGKVALIRGCEPFGDKSVLAAKAAASIYNNVAGSMAGTIGAOSDKG- 173
DB 379 EDFE-KDVAGKIALIRGIDFKDKVAAKAKGAVGLTYDN-----QDKGFP 424
QY 173 -----PYSATYGISLEBQOKLIKLAEGSVSVLDVMSKQENRTTYN--VAOTKGD 223
DB 425 IELPNVDQMPAAFISSKDG-----LLKDNPKQITTFNATPKVLPFTAGT 469
QY 224 PNNVVALGHDPSVAGGINDGSGIISNLVIK--ALTOYSVKNANRFLFTAEFEG 281
DB 470 KLSRFSWGLTADGNIKRDIAPAGODIISVANNKYAKLSGTSKAPL-----VAGING 524
QY 282 LGSNYYVSHLNTLNLRIYLNFDMIASPNYALMITYDGDSAF--NOSGPAGSAQIEK 338
DB 525 LQKQETQYPMTPSERDLAKKVL-----SSATLALYDEDEKAYSPRQOG-AGAVDAK 579
QY 339 LF-----EDYDSDIDLPRIPTQFD-----GRSD-----YEAFLNGISPGGLF 377
DB 580 ASAATMYTDKDNSTSKYHLNNVSDKFEVTVVHNKSDKPOELYQAATVQTDKVDGKHF 639
QY 378 GAEIGIMSE-----ENASRWGQAGVAYDANYHAAGDNMTNLNHEAFILNSKATFA 428
DB 640 LAPKVLYESMOKITIPANSSK---QYVTPIDAS-RFSKDLAOKKNQGYFL--EGFVAF 693
QY 429 VATYANDLSIP 440
DB 694 ODPKKEELMSIP 705
```

RESULT 7

US-08-589-756-1

Sequence 1, Application US/08589756

Patent No. 5846547

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: STREPTOCOCCAL C5a PEPTIDASE VACCINE

NUMBER OF SEQUENCES: 5

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/589,756

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1164 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-589-756-1

Query Match 5.8%; Score 145; DB 2; Length 1164;
 Best Local Similarity 22.8%; Pred. No. 3.9e-05;
 Matches 91; Conservative 43; Mismatches 128; Indels 138; Gaps 18;

QY 1 GRALVSPDEPEDIQLEDLLEGSOQLEDFAYATPERNRV-----FGKA-----H 45
 DB 262 GNALVYANLPDETK-----KPFYAKSKGVRIYTAGNDSSFGKTRRLPLADH 310
 QY 46 DDTVNTLYEELKRTGYDYVKKOPVHLSNADOTLKVGDEIEAKTWTFSVEVTADVA 105
 DB 311 PD-----YGVYGT-----AAADSLTYA-----SYFDNQLT-ETA 341
 QY 106 VVK-----NLGSEADYPSDVEGKVALIKGCEPFGD 137
 DB 342 MKTDDQOQKEMPVLSTNREPKNAYDYAVANRGMKEDF-KDYKRIALIERSDIDFTD 400
 QY 138 KSVLAKAKAAASTIYNNVAGSMAGTIGAAQSDG-----PYSAIVGSLDEGKLI 169
 DB 401 KINAKKAGAVGLIYDN-----QKGFPIELPNVDQMPAEYSRSDG----- 444
 QY 190 KLAEGSVSYDLWVDSKQENRTYV---VAOTKGGDPNNVVALGHTDSEVAGPGIND 246
 DB 444 -----LLKKNQSKQITTFNATPKVLPASGTRKLSRFSWGILTAGNKPDIAP 492
 QY 247 GSGIISNLVIAR--ALTOYVKNNAVRELFETAEFGLGSGNYVSHLATELNKIRLYN 304
 DB 493 GQIILSSANNKRYAKLSGTMSAPL-----VAIVMGLQKQYEQYDPDMQSERLDLAKK 547
 QY 305 FDMIASPNVLMYDGDGSAF---NOSGPRGSAQIEKLE 341
 DB 548 VLM-----SSKATALYDEDEKAFESPROOG-AGAVDAKKAASE 582

RESULT 8
 US-08-547-197-1
 ; Sequence 1, Application US/08547197
 ; Patent No. 5691157
 ; GENERAL INFORMATION:
 ; APPLICANT: Gong, Joseph K.
 ; APPLICANT: Glomski, Chester A.
 ; TITLE OF INVENTION: A METHOD FOR DETECTING A MAMMAL'S PRIOR
 ; TITLE OF INVENTION: EXPOSURE TO RADIATION OR RADIONUCLIDIC AGENTS
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
 ; STREET: Clifton Square, P.O. Box 1051
 ; CITY: Rochester
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 14603
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/547,197
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rogalsky, Peter
 ; REGISTRATION NUMBER: 38,601
 ; REFERENCE/DOCKET NUMBER: 19226/580 (R-5228)
 ; TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 263-1634
 TELEFAX: (716) 263-1600
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 622 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-547-197-1

Query Match 5.6%; Score 140.5; DB 1; Length 622;
 Best Local Similarity 20.4%; Pred. No. 3.8e-05;
 Matches 123; Conservative 82; Mismatches 194; Indels 203; Gaps 35;

QY 14 IQLEDLLEGSOQLEDFAYATPERNRVFGKAHDYVNTLYEELKRTGYD-----YKOP 68
 DB 1 IEFDTII---KQLSQNTYTPRE-----AGSQDENLAYIENL-----FHDKFSKYWRDE 48
 QY 69 ---QVHLSNADQ---TLKVGD--EEIEAKT--MTYSPSEVTADYAVVKNIGC-----SE 114
 DB 49 HYVKIQVKNVSQNLVYINSQNSIDVEAPEGYAFASKAGEYTGKL-VHANFGTKKDFEE 107
 QY 115 ADYPSDVEGKVALIKGCEPFGDGSVLAKAKAAASTIYNN-----VAGSMA 161
 DB 108 LNT---SVNGSLVTRAGKLTFAKAVANAOSEFNIGVLTMDRNTFPVEADLOFGHAHL 165
 QY 162 GT-----LGAOSDKGPYSALVIGISLDEQKLIKLAEG----- 196
 DB 166 GTDDPYTPGPSPFNHTQFPSPQSGLPSIPVQITSAPEKLEKKNEGCPSMNIDSSG 225
 QY 196 -----SVSYDLWVDSKQENRTYVNVVQTKG-GDPNNVVALGHTDSEVAGPGI--NND 246
 DB 226 KIELSONONKVTNNVNLKETRILNIFGVIKGYEEDRYIVGAQDA--WGPVAKSSV 283
 QY 247 GSGIISNLVIARALTOYVKNNAVRELFETAEFGLGSGNYVSHLATELNKIRLYN 294
 DB 284 GTGLL--LKLQVFSMDISKDGFRRPSRIIFASWTAGDYGANVPTWLGISLHLKAF 341
 QY 294 TELNKIRLYL---NFDIASP--NYALM-----TYDG-----DG 322
 DB 342 TYINLKVYVLTGTSNFKVYASPLTYTLMKIMQDVKHPIDGKYLXRNMSMISKIELSLDN 401
 QY 323 SANOSGPRGSAQIEKLF---EDYISIDLPRIPTQFDGRSDYEAT-----LNGLPSCG 374
 DB 402 AAFPLAYSGIPAVSFCFEDEBY-----PYLGTKLD--TYEILQKVPOLNQW---- 449
 QY 375 LFTGAEIGISEENASRWGQAGVADYANHAAGDNNTLNHFAFL-----INSKATAPA 428
 DB 449 VRTAAE-----VAGQFIKLTHTDELTLDYEMTNSKLISPM 484
 QY 429 --VATYANDLSI-----PKRNTTSLHRRART-----MRPGKRAPKTH 466
 DB 485 KDLNCKKADIKDGLSLQWLYSARQGYFRATSLTDFINAEKTNFVAREINDRLMKVE 544
 QY 467 AH 468
 DB 545 YH 546

RESULT 9
 US-08-929-922B-2
 ; Sequence 2, Application US/08929922B
 ; Patent No. 5994113
 ; GENERAL INFORMATION:
 ; APPLICANT: Kauppinen, Sakari
 ; APPLICANT: St. Joan Of
 ; APPLICANT: Spendler, Tina
 ; APPLICANT: Dammann, Claus
 ; APPLICANT: Halkier, Torben
 ; APPLICANT: Stergaard, Peter
 ; APPLICANT: Patkar, Shankant Anant

APPLICANT: Hansen, Kim
TITLE OF INVENTION: An Enzyme With Aminopeptidase Activity
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: No. 59941130 No. 5994113disk of No. 5994113th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/929,922B
FILING DATE: 15-SEP-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4355,204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 377 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-929-922B-2

Query Match 5.4%; Score 134.5; DB 2; Length 377;
Best Local Similarity 23.3%; Pred. No. 6.3e-05;
Matches 66; Conservative 48; Mismatches 98; Indels 71; Gaps 14;

QY 186 QKLLKLAGSVSDLMVDSKOENTTYNVAQTKGDPNNVVALGTTDSVE----- 239
D 134 QEIIISASGAKGVTEVPFKHSPEQS---SLAKIPGKSDKTIY-LGAQDSINLDSPEEG 188
QY 229 AGPGINDSGGIISNLVIAKA-LIOTSY-----KNAVRFLEMTAEFFGLSNNYVSLN 292
D 189 RAPGADDDGSGVVTLEFRVLIDEXVAGAPVTFEHHYAGEGGLDSQ-DIFEOY 247
QY 293 ATELKIRLYLNFDM-----IASPNYALMIYDGDGSAFNOSGPAASQIEKLFEDY 343
D 248 SOKSDVAKAMLQODWTGKGTDTAGKPEISGIITD-----NVDENLTKFKLVYDA 299
QY 344 YDSIDLPIHPOFGRSYEAFLINGIFSGGLFTGABGIMSEENASRKGGAGVAYDANY 403
D 300 YCT-----IPT-VDSKCY-----GCSDHASATKYGYPAAAFESAF 335
QY 404 -----HAAGDMNTLNHEAFLINSKAT---AFAVATYANDL 436
D 336 GDDSPYIHSADDTIETVNFHYVHGLTLGFAVELA-FAOSL 377

RESULT 10
US-08-547-197-2
Sequence 2, Application US/08547197
Patent No. 5691157
GENERAL INFORMATION:
APPLICANT: Gong, Joseph K.
APPLICANT: Giontski, Chester A.
TITLE OF INVENTION: A METHOD FOR DETECTING A MAMMAL'S PRIOR
TITLE OF INVENTION: EXPOSURE TO RADIATION OR RADIONUCLIDIC AGENTS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051

CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/547,197
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rogalsky, Peter
REGISTRATION NUMBER: 38,601
REFERENCE/DOCKET NUMBER: 19226/580 (R-5228)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1634
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 760 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-547-197-2

Query Match 5.2%; Score 130; DB 1; Length 760;
Best Local Similarity 20.1%; Pred. No. 0.00056;
Matches 72; Conservative 58; Mismatches 130; Indels 98; Gaps 15;

QY 92 MTSPSVETADVAVNKLGSE--ADYPSDVEKVALIRGCEPGEKSVLAKAKAA 149
D 220 VAYSKAATVTKL-VHANFGTKKDFEDLYTPVNSIYVBAKITFEKVA NESLVAIG 278
QY 150 SIYNN-----VAGSMAGTIGAASDKG-PYS-----AIVGIS 181
D 279 VLIYMDQTKPIYNAELSPFGHAHGTGDPYTPGFPSFNHTQPPSSSGLPINIVOTIS 338
QY 182 LEDQKLI-----KLAEGSVSDLMVDSKOENTTYNVAQTKG-CD 223
D 339 RAAEKXLFGNMGGCPDMDKTDSTCRVYTSSEKNVKTIVSNVLKEIKILNIFGVIGFE 398
QY 224 PNNVVALGHTDSYEAQPGINDSGGIISNLVIAKALTQYSVKN-----AVRFLEWTAE 277
D 399 PDHYVVVGAQRDA--WCPGAKSGVGTALLKLAQMSDVLKDGFOPSRSIIFASVSAG 456
QY 278 EFGLLGSNNYS-----HNA-TELKIRLYL--NEDMIASPNYALMIYDGDGSAFNQ 327
D 457 DFGSVGATEWLEGYLSLHKAFTYINLDAVAGTSNFKVSASPLTYTL---EKTQNV 513
QY 328 SGPA-----GSAQIEKLFED-----YDSIDLPIHPTOPD 357
D 514 KHPYTGQFIVODSNWAKVEXKLIDNAPFLAYSGIPAVSFCEBETDYPYIGTMD 571

RESULT 11
US-08-542-003-6
Sequence 6, Application US/08542003
Patent No. 5864013
GENERAL INFORMATION:
APPLICANT: Goldberg, Edward B.
TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
TITLE OF INVENTION: NANOMETER STRUCTURES AND USE THEREOF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie and Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York

```

      COUNTRY: US
      ZIP: 10036
      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Patent Release #1.0, Version #1.25
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/542,003
      FILING DATE: 13-OCT-1995
      CLASSIFICATION: 530
      ATTORNEY/AGENT INFORMATION:
      NAME: Mistrock, S. Leslie
      REGISTRATION NUMBER: 18,872
      REFERENCE/DOCKET NUMBER: 8471-0005-999
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (212) 790-9090
      TELEFAX: 212-869-8864
      TELEX: 66441 PENNTE
      INFORMATION FOR SEQ ID NO: 6:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 1026 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      ORGANISM: Bacteriophage T4
      IMMEDIATE SOURCE:
      CLONE: p37 amino acid
      US-08-542-003-6

Query Match      4.4%; Score 109; DB 2; Length 1026;
Best Local Similarity 18.2%; Pred. No. 0.1;
Matches 112; Conservative 69; Mismatches 164; Indels 272; Gaps 31;

```

```

      COUNTRY: US
      ZIP: 10036
      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Patent Release #1.0, Version #1.25
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/322,760A
      FILING DATE: 13-OCT-1994
      CLASSIFICATION: 530
      ATTORNEY/AGENT INFORMATION:
      NAME: Mistrock, S. Leslie
      REGISTRATION NUMBER: 18,872
      REFERENCE/DOCKET NUMBER: 8471-0003-999
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (212) 790-9090
      TELEFAX: 212-869-8864
      TELEX: 66441 PENNTE
      INFORMATION FOR SEQ ID NO: 6:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 1026 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      ORGANISM: Bacteriophage T4
      IMMEDIATE SOURCE:
      CLONE: p37 amino acid
      US-08-322-760A-6

Query Match      4.4%; Score 109; DB 2; Length 1026;
Best Local Similarity 18.2%; Pred. No. 0.1;
Matches 112; Conservative 69; Mismatches 164; Indels 272; Gaps 31;

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OK protein - protein search, using sw model

Run on: May 10, 2000, 19:08:05 ; Search time 47.2 Seconds
(without alignments)
597.477 Million cell updates/sec

Title: US-09-080-127-2_COPY_16_496
Perfect score: 2500
Sequence: 1 GRALVSPDERPEDIDLEDL.....APKTHAVSGSGVHSCVEA 481

Scoring table:
BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 168808 segs, 58629743 residues

Total number of hits satisfying chosen parameters: 168808

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : PIR_63:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	782.5	31.3	537	2	A54134
2	668	26.7	500	2	H70629
3	467.5	18.7	284	2	S66427
4	317	12.7	455	2	S39663
5	189	7.6	368	2	B70463
6	188.5	7.5	806	2	A41341
7	163	6.5	751	2	T30154
8	160	6.4	1902	2	S06997
9	159	6.4	1902	2	B45764
10	158	6.3	565	2	E71223
11	157	6.3	374	2	S69699
12	156.5	6.3	504	2	S24314
13	156	6.2	1902	1	B44856
14	154.5	6.2	1483	2	T19751
15	154	6.2	750	2	A56881
16	152	6.1	413	2	A54467
17	151	6.0	1962	2	A32634
18	149.5	6.0	561	2	F75191
19	149.5	6.0	757	2	A48592
20	149	6.0	763	1	S29548
21	149	6.0	1167	1	A35066
22	148	5.9	776	1	JH0570
23	140.5	5.6	622	2	A34549
24	138.5	5.5	591	2	A75474
25	136	5.4	794	2	T40289
26	133.5	5.3	811	2	S57149
27	130	5.2	760	1	UXHU
28	129	5.2	925	2	S69539
29	119	4.8	654	2	B75587
30	116.5	4.7	464	2	C69732

31	114.5	4.6	1770	2	A71517	hypothetical prote
32	112	4.5	380	2	T17236	hypothetical prote
33	111.5	4.5	325	2	E64998	hypothetical prote
34	109.5	4.4	345	1	K2EC	alkaline phosphata
35	109.5	4.4	567	2	C72698	hypothetical prote
36	109	4.4	1026	1	TUBP74	tail fiber protein
37	109	4.4	1148	2	S72635	exo-poly-alpha-gal
38	108.5	4.3	404	2	A44833	serine proteinase
39	108.5	4.3	2334	2	S32920	cell wall-associat
40	108	4.3	428	2	C70989	hypothetical prote
41	108	4.3	504	2	S33189	phase-1 flagellin
42	107	4.3	504	2	S33193	phase-1 flagellin
43	107	4.3	690	2	C69208	DNA helicase relat
44	107	4.3	1705	2	S51672	adenylate cyclase
45	106.5	4.3	839	2	F64171	hypothetical prote

ALIGNMENTS

RESULT 1

A54134

aminopeptidase Y (EC 3.4.11.-) precursor, vacuolar - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YBR2024; protein YBR286w

C:Species: Saccharomyces cerevisiae

C>Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 29-Oct-1999

C:Accession: A54134; S44548; S45527; A54133; S39142

R:Nishizawa, M.; Yasuhara, T.; Nakai, T.; Fujiki, Y.; Ohashi, A.

J. Biol. Chem. 269, 13651-13655, 1994

A:Title: Molecular cloning of the aminopeptidase Y gene of Saccharomyces cerevisiae.

A:Reference number: A54134; MUID:94230479

A:Accession: A54134

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-537 <NIS>

A:Cross-references: GB:L11635; NID:9469463; PIDN:AA19559.1; PID:9469464

R:Holmstrom, K.; Brandt, T.; Kallensee, T.

Yeast 10(Suppl. A), S47-S62, 1994

A:Title: The sequence of a 32420 bp segment located on the right arm of chromosome II

A:Reference number: S44537; MUID:94378722

A:Accession: S44548

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 'MDGNTMFSLQQLVEARKKKTKTPT', 1-13, 'LR', 16-537 <HOL>

A:Cross-references: EMBL:X76033

A>Note: this sequence has been revised in reference A54134

A>Note: It is assumed that the second ATG codon within the open reading frame is the R.Brandt, T.; Christiansen, C.; Holmstrom, K.; Kallensee, T.

A:Reference number: S46157

A:Accession: S46158

A:Molecule type: DNA

A:Residues: 'MDGNTMFSLQQLVEARKKKTKTPT', 1-13, 'LR', 16-537 <BRA>

A:Cross-references: EMBL:Z36155; MIPS:YBR286w

A>Note: this sequence has been revised in reference A54134

A>Note: It is assumed that the second ATG codon within the open reading frame is the R.Nishizawa, M.; Yasuhara, T.; Nakai, T.; Fujiki, Y.; Ohashi, A.

A:Title: Molecular cloning of the aminopeptidase Y gene of Saccharomyces cerevisiae

A:Reference number: S45527

A:Accession: S45527

A:Molecule type: DNA

A:Residues: 1-537 <NIK>

A:Cross-references: EMBL:L31635; NID:9469463; PIDN:AA19559.1; PID:9469464

R:Yasuhara, T.; Nakai, T.; Ohashi, A.

J. Biol. Chem. 269, 13644-13650, 1994

A:Title: Aminopeptidase Y, a new aminopeptidase from Saccharomyces cerevisiae

A:Reference number: A54133; MUID:94230478

A:Accession: A54133

A:Molecule type: protein

A:Residues: 57-70 <YAS>

A>Note: 70K and 75K forms had the same amino-terminal sequence and appeared to differ C:Genetics:

RESULT **4**

S39663
aminopeptidase homolog ywad - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 07-Oct-1994 #sequence_revision 24-Feb-1995 #text_change 15-Oct-1999
A:Accession: S39663; S16427; D70050
R:Glaser, P.; Kunst, F.; Arnaut, M.; Couderc, M.P.; Gonzales, W.; Hullo, M.F.; Ionescu,
A.; Rapoport, G.; Danchin, A.
Mol Microbiol 10, 371-384, 1993
A>Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region fr
A:Reference number: S39655; MUID:95020537
A:Accession: S39663
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-455 <GLA>
A:Cross-references: EMBL:X73124; NID:g413923; PIDN:CAA51564.1; PID:g413932
R:Glaser, P.; Kunst, F.; Debarbouille, M.; Vertes, A.; Danchin, A.; Dedonder, R.
DNA Seq. 1, 251-261, 1991
A>Title: A gene encoding a tyrosine tRNA synthetase is located near sacs in Bacillus sub
A:Reference number: S16421; MUID:92216127
A:Accession: S16427
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-68 <GL2>
A:Cross-references: EMBL:X52480; NID:g40236; PIDN:CAA63725.1; PID:g40243
R:Kunst, F.; Ogasawara, N.; Mosser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berthelette,
C.; Brog, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capriano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 350, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funo, S.; Gallizi, A.; Gallie
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigsberg, L.G.; Krogh, S.; Kumano, M.; Kurita, K.; Ladakis, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee
Y. M.; Ogawa, K.; Ogikawa, A.; Outega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Potte
Rieger, M.; Rivolta, C.; Rocho, E.; Rocho, B.; Rose, M.; Sadai, Y.; Satou, T.; Scanon
A:Authors: Schleich, S.; Schroeder, R.; Scoffone, F.; Sekiguchi, J.; Sekowa, A.; Seroc
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpiltra, P.; Togonni, A.; Tosato, V.; Uchiyama
T.; Winiers, P.; Wipet, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A:Authors: Yoshikawa, H.F.; Zumsstein, E.; Yoshikawa, H.; Danchin, A.
A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:96044033
A:Accession: D70050
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-455 <KUN>
A:Cross-references: GB:299123; GB:AL009126; NID:g2636240; PIDN:CAB15873.1; PID:e1186346
A:Experimental source: strain 168
C:Genetics:
A:Gene: ywad

Query Match 12.7% Score 317; DB 2; Length 455;
Best Local Similarity 27.4% Pred. No. 3e-14;
Matches 119; Conservative 67; Mismatches 174; Indels 74; Gaps 16;

OY 19 LLEGSQQLDEFAYAY-----PERRRVFGGKAHDDTVNYLYEEL-----KKTGY--- 62
|| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 18 LLRPAGSYTPRAHANVOISNSRELFPRKAKHAYSTISQSLSAEGPRIAGTAARAKKALLIA 77
OY 62 -----YDVYKQPQVHLMSNADQTLKVGDDEIEIEMKMTYSPSEVTADVAVYKKLGCS- 114
|| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 78 SSMRKLIKLDV-KVGPFNIIPDRLEGTLSSAGRDIILQAASGSAPEEOGLAPILYNNAIGLY 136
OY 114 ENDYSVDVEGRKALKRGCGFRGDSYLAARAKAASTYYNNVASMGATLGAAOSDGP 173
|| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 137 QKDFTPADAKGTALLSRGDTTYERAKKNNEAGAARAIVIIYNKESLVPTPMISGNKVG- 196
OY 174 YSAIYGISLEDOKLIKLTAAGSVSVDLTWVDSKOENRTTYNNVAOTKGDDPNK-----VV 228
|| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 196 -IPVGIKKEDEALTOQGEA-TLTKAKAF-----NQTSQNITIGKK--PKNIKHPDIV 245
OY 229 ALGGHTDVSVEAGPGINDSGSIISNLVIATAKALTQYSVKNAVFELFWLAEFGILGCSNTYV 288

```

D6      246      YTTAAHYADVPSPGANDNGSGTAVLEAKVRLKVSPPDKREIRTFAGAEELGSSSHYV 305
              :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY      289      SHLNATLEAKRLRYLNFPMI-----ASPNTALMIYDGDGSAFNQSPAGSAQLEKLFED 342
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db       306      DHLSEKLEKRESEVENFNDLVGTSEKASSELVNTL-DGQSNYWESSRTAEKIG----- 360
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY      343      YVSDIDLPHTPEFGDRSDYEAFLNGIPSGGLFTGAEGIMSEENASRMWGQACVAYDAN 402
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db       360      -FDLSLST-----TQ-GGSSDHPHFHEAGIDSANFT-----WGDPETEEYEPW 399
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY      403      YHAAGDNMTNLNHE 416
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db       400      YHTPEDSIEHISKE 413
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT      5
B70463
hypothetical protein ag_1891 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence.revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: B70463
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:9819666
A:Accession: B70463
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1368 <AOP>
A:Cross-references: GB:AE000762; NID:g2984163; PIDN:AAOC07705.1; PID:g2984172; GB:AE000
A:Experimental source: strain VFS
C:Genetics:
A:Gene: ag_1891

Query Match      7.68; Score 189; Db 2; Length 368;
Best Local Similarity 23.48; Pred. No. 1.2e-05;
Matches 79; Conservative 66; Mismatches 147; Indels 46; Gaps 13;

OY      36      RNRYVFGKAHDDTVLYLEELKTKGYDYVYKQP-QVHLMSNADOTLKVGDEIEAKTMTY 94
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db       19      KNRLSGTGNKEARFKINYLKRG-FKVEDEDSVTKTPVVSARICEGEPEFAPFLIG 77
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY      95      SPSVEVTLADVAVVKNLGCSEADYPSDVEGKALIKRGCCPFGDKSVLAAKKAAASIYVN 154
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db       78      SLMGSTSEVAILKRL-----EEEKDLKGKVAAPVGGRRDSEAKFLRDKAKAGMYTF- 132
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY      155      NVAGSMAGTLGAAGDKPQSAIVG-----ISLEDGOKILIKAEAGSVSDLWVSKQ-EN 209
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db       132      -----LELDVVPFSGTIGDVKFPFVAVNTTRELVAIKIEGKNVILDIRTEKEIKG 179
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY      210      RTTYVVAQTRGQDPNNVVALGHTDVEAGPAGINDDSGIIISNLVIAKALTYQ-SYKNA 268
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db       180      KRLV--VEFGRG---PYLLVAHYDTKPFYGAIDNGLSVALLVMSRELAGEQEIFPR 233
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY      269      VAFLEWTFEEFGGLGSNTYVSHLNATELNKRLILNPFMISPYVALMIYGGDSAFNQ 328
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db       234      IRLTDEEELGEGASHALH-----SKNFYVINLDSIGMKPNPALYEDA-----S 281
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY      329      GPAGSAQLEKLFEDYYD-SIDLPHITQFDRSDYEAF 365
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db       282      GENGRLINELKFLHLMDKVDIPFVESK-TGMSDHPF 318
              :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT      6
A11341
microbial serine proteinase (EC 3.4.21.-), minor (vpr), precursor - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 17-Jul-1992 #sequence.revision 17-Jul-1992 #text_change 29-Sep-1999
C:Accession: A11341; B11341; S39700; D6930
R:Sloma, A.; Rifo Jr., G.A.; Theriault, K.A.; Dwyer, M.; Wilson, S.W.; Petro, J.

```


QY 422 -SKATAFAVATYANDLSIPKRTSSLRHAPTRPFGKAPKTAHANVSGSCWHSQVE 480
DB 567 IGYWMAELAKTEFADYV--LIPNMTT---HFASVMLKTY--LPQLTITTSIGINVSRSDFE 618

RESULT 8

serine proteinase (EC 3.4.21.-) precursor - Lactococcus lactis plasmid pLP763
C:Species: Lactococcus lactis
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 22-Jun-1999
C:Accession: S06997; S08082
R:Kimaki, M.; Ikemura, H.; Shimizu-Kadota, M.; Hatahima, A.
Mol. Microbiol. 3, 359-369, 1989
A:Title: Molecular characterization of a cell wall-associated proteinase gene from Streptococcus lactis
A:Reference number: S06997; MUID:89313288
A:Accession: S06997
A:Molecule type: DNA
A:Residues: 1-1902 <KIM>
A:Cross-references: EMBL:X14130; NID:g47197; PIDN:CA43250.1; PID:g47199
C:Genetics:
A:Genome: plasmid pLP763
C:Superfamily: cell envelope-associated serine proteinase; subtilisin homology
C:Keywords: hydrolase; serine proteinase; transmembrane protein
F:208-634/Domain: subtilisin homology #status atypical <SBT>

Query Match 6.4%; Score 160; DB 2; Length 1902;
Best Local Similarity 21.7%; Pred. No. 0.012;
Matches 109; Conservative 57; Mismatches 152; Indels 184; Gaps 24;

QY 80 LKVGDEIEAKMTYPSVSE-----VTADVAVYKNGCSEADYPSDEKVALIKRGEC 134
DB 440 LQGPETIQSSNDFTGSFDOKKFFVVKDASGNLSKG-KVADITADPKGKIAIVKRGELS 498
QY 135 FGDKSVLAARAKAASIVYNNVAGSMAGTGAASDKGYSIAVIGSLDEGOKLIKLAEA 194
DB 499 FADKQYVAAQAAGAGLIIYNN-DGTAIPVTSMAITTFP--TFGLISVTGOKLVDWVTA 554
QY 195 ---GSYSVDL-----WVDSKQENRTTYNNVAQTGKGDNNVVALGGHDSVEAGPG 242
DB 555 HPDDSLGVKIALTLVFNOKYTEDKMSDFTSYGVSNL--SFKPDITAPGGINWSTQNNNG 612
QY 243 -----IN 244
DB 613 YTNMGSTNASPFIAQSALLKQALNNKNPFYAYYKQKGTALDTFLKTVEMNTAQPIN 672
QY 245 D-----DGSGLISNLYIAKALQYSYKN-----AVRFLFTAE--F 279
DB 673 DINYNNVIVSPRQAG--LVYKAAIDALEKNPSTVAENGIVAVELKDTSTDKTF 728
QY 280 GLIGSVYVSHL-----NATELNKIRLYLNFDMIASPNALMIDG--DGSAF--NOS 328
DB 729 KLFTNRTTHLTYQMDSTDTNAV-----YTSATDPNSG-VLYDKKIDGAIAIKGSNT 782
QY 329 GPAG-SAQIEKLFEDYDSIDLPHIPTQDGRSDYAFI----- 367
DB 783 VPAGKTAQIEFTLS-----LPKSPDOQFEGEGLNFKSGDSRLNLPYMGFFGDM 832
QY 367 -----LNGI-----PSGGLFTGAEIGSEENASRMGQAGVAYDANHYAAGDNMTNLN 415
DB 833 NGKIYDSLNGITISPAAGNF-GTVPLLTNNK-----TGTYGGMVTDADGNGQTVDD 884
QY 416 EAFILNSKATAFAVATYANDLS 437
DB 885 QAIASFSSDKNAL-----YNDIS 901

RESULT 9

B45764
serine proteinase (EC 3.4.21.-) precursor, cell-envelope-associated - Lactococcus lactis
C:Species: Lactococcus lactis subsp. cremoris
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Mar-1998

C:Accession: B45764; A60460
R:Kok, J.; Leenhouts, K.J.; Haandrikman, A.J.; Ledebwer, A.M.; Venema, G.
Appl. Environ. Microbiol. 54, 231-238, 1988
A:Title: Nucleotide sequence of the cell wall proteinase gene of Streptococcus cremor
A:Reference number: A45764; MUID:88149035
A:Accession: B45764
A:Molecule type: DNA

A:Residues: 1-1902 <KOK>
A:Cross-references: GB:M24767
R:laan, H.; Konings, W.N.
Appl. Environ. Microbiol. 55, 3101-3106, 1989
A:Title: Mechanism of proteinase release from Lactococcus lactis subsp. cremoris Wg2.
A:Reference number: A60460
A:Accession: A60460
A:Molecule type: protein
A:Residues: 188-193 <LA>
C:Comment: This proteinase is proposed to exist in both a cell-wall-associated and an
C:Genetics:
A:Genome: plasmid
C:Superfamily: cell envelope-associated serine proteinase; subtilisin homology
C:Keywords: extracellular protein; hydrolase; membrane bound; serine proteinase
F:1-33/Domain: signal sequence #status predicted <SIG>
F:134-187/Domain: propeptide #status predicted <PRO>
F:188-1902/Product: serine proteinase; cell-envelope-associated #status experimental
F:208-634/Domain: subtilisin homology #status atypical <SBT>

Query Match 6.4%; Score 159; DB 2; Length 1902;
Best Local Similarity 21.7%; Pred. No. 0.015;
Matches 109; Conservative 58; Mismatches 151; Indels 184; Gaps 24;

QY 80 LKVGDEIEAKMTYPSVSE-----VTADVAVYKNGCSEADYPSDEKVALIKRGEC 134
DB 440 LQGPETIQSSNDFTGSFDOKKFFVVKDASGNLSKG-ALADYTADAKGKIAIVKRGELS 498
QY 135 FGDKSVLAARAKAASIVYNNVAGSMAGTGAASDKGYSIAVIGSLDEGOKLIKLAEA 194
DB 499 FADKQYVAAQAAGAGLIIYNN-DGTAIPVTSMAITTFP--TFGLISVTGOKLVDWVTA 554
QY 195 ---GSYSVDL-----WVDSKQENRTTYNNVAQTGKGDNNVVALGGHDSVEAGPG 242
DB 555 HPDDSLGVKIALTLVFNOKYTEDKMSDFTSYGVSNL--SFKPDITAPGGINWSTQNNNG 612
QY 243 -----IN 244
DB 613 YTNMGSTNASPFIAQSALLKQALNNKNPFYAYYKQKGTALDTFLKTVEMNTAQPIN 672
QY 245 D-----DGSGLISNLYIAKALQYSYKN-----AVRFLFTAE--F 279
DB 673 DINYNNVIVSPRQAG--LVYKAAIDALEKNPSTVAENGIVAVELKDTSTDKTF 728
QY 280 GLIGSVYVSHL-----NATELNKIRLYLNFDMIASPNALMIDG--DGSAF--NOS 328
DB 729 KLFTNRTTHLTYQMDSTDTNAV-----YTSATDPNSG-VLYDKKIDGAIAIKGSNT 782
QY 329 GPAG-SAQIEKLFEDYDSIDLPHIPTQDGRSDYAFI----- 367
DB 783 VPAGKTAQIEFTLS-----LPKSPDOQFEGEGLNFKSGDSRLNLPYMGFFGDM 832
QY 367 -----LNGI-----PSGGLFTGAEIGSEENASRMGQAGVAYDANHYAAGDNMTNLN 415
DB 833 NGKIYDSLNGITISPAAGNF-GTVPLLTNNK-----TGTYGGMVTDADGNGQTVDD 884
QY 416 EAFILNSKATAFAVATYANDLS 437
DB 885 QAIASFSSDKNAL-----YNDIS 901

RESULT 10

B45764
hypothetical protein PH0048 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii


```
Db 253 NTDKMPSTVQMGSVYIGLSDPRTPAPPSIGDLFEKTEQDLDEKIPITPMLPTVAT 312
QY 185 GQKLIKLAEGSVSD-----LWDSKQENNT----- 213
Db 313 AQLFFENKKGAVNADFGKLNVTYRYGPGLINNOKLVTVAENEERSVIVELISRLR 372
QY 213 -----YVNVQATGGG-DPNNVVALGHTDSVEAGPGINDGSGGISNLVIKALTOXSV 265
Db 373 LNFRIQINIMGYIKSGQEPDFVLVSNHYDAWTY--GAVDPNSGISTILEVSRALKQTON 430
QY 266 K-----NAVFLEFWTAEFFGLSGSYVSHLNATELKRILYNFDMI-----AS 310
Db 431 QTGMIPARSILFAHWDADEYGLISTEFAEYRQLMRRAVAVIMDLIGNQTLGLSN 490
QY 311 PNYALMIDGGSAPFNOGSPGASQIEKLFEDYDSIDLPPIPTOFDGRSDYEAFLNGI 370
Db 491 PTVANVL---RSAANVNEOP-NPTEMEOGRKTLXDSMKY-YAPSK-NNRSTHP--YORI 541
QY 371 PSGG 374
Db 542 PAGG 545
```

RESULT 15

```
A56881
Prostate-specific membrane antigen - human
C:Species: Homo sapiens (man)
C:Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 05-Nov-1999
C:Accession: A56881
R:Izraeli, R.S.; Powell, C.T.; Falt, W.R.; Heston, W.D.
Cancer Res. 53, 227-230, 1993
A:Title: Molecular cloning of a complementary DNA encoding a prostate-specific membrane
A:Reference number: A56881; MUID:93113576
A:Accession: A56881
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-750 <ISR>
A:Cross-references: GB:M99487; NID:g190663; PIDN:AAA60209.1; PID:g190664
A:Experimental source: prostatic carcinoma cell line LNCaP
A:Note: sequence extracted from NCBI Backbone (NCBIN:121724, NCBIPI:121725)
C:Keywords: surface antigen; transmembrane protein
```

```
Query Match 6.2% Score 154; DB 2; Length 750;
Best Local Similarity 20.1% Pred. No. 0.0084;
Matches 90; Conservative 59; Mismatches 170; Indels 128; Gaps 17;
```

```
QY 28 DFAYAYPERNRVFGKAHDDTVNTLYE--ELKRTGYDYVKQPOVHLMSNADQTLKAYGD 84
Db 114 DVLSTYP-----NKTHPNYISIIINEDGNEIFNTSLF---EPPPPGYENV----- 155
QY 85 EELIAKTTTYSPEVETADVAVY-----KNLGCSEADYPSDEGVKALIKRGCEPFGDKS 139
Db 155 SDIVPPPSAFSPQGMPEGDLVYVNTARTEDFKLERDKKINCSGKIYIARYGKVFGRNKV 214
QY 140 VLAAKAKAAASIVYNNVA-----GSMAGTL----- 165
Db 215 KNAQLAGKAGYILSDPADYFAPGVKSPDGMNLPGGGVQKGNITILNMGADPLTPGYPA 274
QY 165 -----GAQSDKGPYSALVIGISLEDGOKLIK-----L 191
Db 275 NEYAYRGIADAVGLPSIPVHPIGYDAOKLLEKKGSAAPPDSSWRGSLKVPYVNGPGFT 334
QY 192 AEAGSVSYDLWVDSKQENRTYNNVAQTKGG-DPNNVVALGHTDSVEAGPGIN-DDGSG 249
Db 335 GNFSTQKVKMHIHSTNEVTRILYVIGTLRGAVPEPRYVILGHRDSWVFG-GIDPQSGAA 393
QY 250 IISNLVIKALTO--XSVKNAVRELFWTAEFFGLSGSYVSHLNATELKRILYNFND 306
Db 394 VVHEIVTSFGTLKEGMRPRRTILFASWDAAEFGLGSTWAAEKSRLQERGAAYINAD 453
QY 307 MIASPNYALMIDGGSAPFNOGSPGASQIEKLFEDYDSIDLPPIPTOF 356
```

```
Db 454 SIEGNTTLRY---DCTPLWYSLVHNLTKELKSPDEGEFGKSLYESWTKKSPSP-----EF 506
QY 357 DGR-----SDYEAFLN-GIPSG 373
Db 507 SGMPRISKLSGNDFEVFFQRLGIASG 533
```

Search completed: May 10, 2000, 19:34:53
Job time: 1608 sec

Thu May 11 10:03:18 2000

us-09-080-127-2_copy_16_496.rpr

Page 9

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2000, 19:33:12 ; Search time 42.29 Seconds

(without alignments)
346.390 Million cell updates/sec

Title: US-09-080-127-2_COPY_16_496

Perfect score: 2500
Sequence: 1 GRALVSPDFPFDIOLEDL.....APKTHAVSGSGCHMSOVERA 481

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 83857 seqs, 30454973 residues

Total number of hits satisfying chosen parameters: 83857

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_38.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	782.5	31.3	537	1 APE3_YEAST	P37302 saccharomyc
2	467.5	18.7	284	1 APX_STRGR	P80561 streptomyc
3	317	12.7	455	1 YWAD_BACSU	P25182 bacillus su
4	188.5	7.5	806	1 SUBV_BACSU	P29141 bacillus su
5	160	6.4	1902	1 P2P_LACLA	P15293 lactococcu
6	156.5	6.3	504	1 AMPX_VIBPR	Q01693 vibrio prot
7	136	6.2	1902	1 P2P_LACPA	Q03470 lactobacill
8	154	6.2	750	1 PSM_HUMAN	Q04609 homo sapien
9	152	6.1	413	1 YBS4_YEAST	P38244 saccharomyc
10	152	6.1	1902	1 P1P_LACLC	P15292 lactococcu
11	151	6.0	1902	1 P3P_LACLC	P15292 lactococcu
12	149	6.0	1167	1 SCRA_STRPY	P15926 streptococ
13	133.5	5.3	811	1 YJ96_YEAST	P47161 saccharomyc
14	130	5.2	760	1 TRSR_HUMAN	P02786 homo sapien
15	129	5.2	925	1 VPH_BPHP1	P51735 bacterioph
16	116.5	4.7	464	1 XKDK_BACSU	P54331 bacillus su
17	111.5	4.5	323	1 YFBL_ECOLI	P16482 escherichia
18	109.5	4.4	345	1 IAP_ECOLI	P10423 escherichia
19	109	4.4	1026	1 VG37_BPT4	P03744 bacterioph
20	108.5	4.3	2334	1 WAPA_BACSU	Q00783 bacillus su
21	108	4.3	504	1 FLIC_SALEN	Q06982 salmonella
22	107	4.3	504	1 FLIC_SALRO	Q06982 salmonella
23	106.5	4.3	839	1 YJDB_HABIN	P43182 haemophilus
24	106	4.2	1705	1 CYAA_BOBR	O57506 bordetella
25	105.5	4.2	427	1 PYRC_LACLE	P43785 lactobacill
26	105	4.2	647	1 CN16_ECOLI	P03749 escherichia
27	104.5	4.2	1132	1 VHSJ_LAMB	P03749 escherichia
28	102.5	4.1	1018	1 EUTB_ECOLI	P13635 bacterioph
29	102.5	4.1	1018	1 ENBA_STAAU	P13788 staphylococ
30	102.5	4.1	1545	1 IGA3_HABIN	P43385 haemophilus
31	102	4.1	504	1 FLIC_SALDU	Q06971 salmonella
32	102	4.1	504	1 FLIC_SALNA	O52959 salmonella
33	102	4.1	1295	1 PURL_ECOLI	P15254 escherichia
34	102	4.1	1540	1 YDBA_ECOLI	P33666 escherichia

ALIGNMENTS

RESULT	1	STANDARD	PRT	537 AA.
APX3_YEAST				
ID	APX3_YEAST			
AC	P37302; P38154;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	AMINOPEPTIDASE Y PRECURSOR (EC 3.4.11.-).			
GN	APX3 OR YBR286W OR YBR2024.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
OC	Saccharomycetaceae; Saccharomycetes.			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 57-70.			
RC	STRAIN-D273-10B.			
RX	MEDLINE: 94230479.			
RA	Nishizawa M., Yasunara T., Nakai T., Fujiki Y., Ohashi A.;			
RT	Molecular cloning of the aminopeptidase Y gene of Saccharomyces			
RT	cerevisiae. Sequence analysis and gene disruption of a new			
RT	aminopeptidase.";			
RL	J. Biol. Chem. 269:13651-13655(1994).			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-S286C;			
RX	MEDLINE: 94378722.			
RA	Holmstrom K., Brandt T., Kallensee T.;			
RT	"The sequence of a 32,420 bp segment located on the right arm of			
RT	chromosome II from Saccharomyces cerevisiae.";			
RL	Yeast 10:847-862(1994).			
CC	-1- COFACTOR: BINDS ONE ZINC ION PER MOLECULE.			
CC	-1- SUBUNIT: MONOMER.			
CC	-1- SUBCELLULAR LOCATION: LYSOSOME-LIKE VACUOLES.			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M33; ALSO KNOWN AS THE			
CC	AMINOPEPTIDASE Y FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sid.ch).			
CC	-----			
DR	EMBL: L31635; AAA19559.1; -			
DR	EMBL: X76053; CAAS3649.1; -			
DR	EMBL: Z36155; CAAS251.1; -			
DR	PIR: S44548; S44548.			
DR	PIR: S39142; S39142.			
DR	HSSP: P80561; IXD0.			
DR	SSD: L00000933; APX3.			
KW	Hydrolase; Aminopeptidase; Zymogen; Zinc; Metalloprotease;			
KW	Glycoprotein; Signal.			
FT	SIGNAL	1	21	POTENTIAL.
FT	PROPEP	22	56	POTENTIAL.
FT	CHAIN	57	537	AMINOPEPTIDASE Y.
FT	CARBOHYD	85	85	POTENTIAL.
FT	CARBOHYD	96	96	POTENTIAL.

35	101.5	4.1	1051	1 CARR_SUTSO	O59969 sulfolobus
36	101	4.0	504	1 FLIC_SALMO	Q06973 salmonella
37	101	4.0	718	1 FLGE_HELPY	P50610 helicobacte
38	101	4.0	810	1 COAT_JCDNV	Q90053 junonia coe
39	101	4.0	1049	1 RT31_ACTPL	P51330 actinobacil
40	100.5	4.0	409	1 AMB2_BACST	O53389 bacillus st
41	100	4.0	714	1 CDG1_PAPMA	P04830 paenibacill
42	99.5	4.0	409	1 AMB1_BACST	P37113 bacillus st
43	99.5	4.0	507	1 FLIC_SALON	Q06974 salmonella
44	99	4.0	371	1 YJDE_BACSU	P54542 bacillus su
45	99	4.0	504	1 FLIC_SALMC	Q06981 salmonella

FT CARBOHYD 115 115 POTENTIAL.
FT CARBOHYD 150 150 POTENTIAL.
FT CARBOHYD 162 162 POTENTIAL.
FT CARBOHYD 371 371 POTENTIAL.
FT CARBOHYD 427 427 POTENTIAL.
FT CARBOHYD 480 480 POTENTIAL.
FT CONFLICT 14 15 POTENTIAL.
SO SEQUENCE 537 AA: 60137 MW: B34B7619194C7F3 CRC64:

Query Match 31.3%; Score 782.5; DB 1; Length 537;
Best Local Similarity 40.6%; Pred. No. 2,4e-43;
Matches 187; Conservative 74; Mismatches 161; Indels 39; Gaps 13;

OY 5 VSPDEPDIEDLLESGQLEDFR-----YAIPEKRVFGKAHDDTYNL---YEE 55
DB 70 VESERKODIKVDLNDATMDYRLANSTPDYGHF--TRVIGSKGHNTMEYILNVEDD 127
OY 56 LKKTGYDYVKQPOYHWMNADQTLKVGDEIE--AKTMT--SPSVE-VTADVAVKN 109
DB 128 MO--DIYDVSLO-EFEALSGLKISFENLSDAETGKSFANTTAALSPVDGFKLVEIPN 184
OY 110 LGCSBADYPSDV-----EGKVALIKRGCEPFGDKSYLAKAARAASTVNNVAGSMAGTL 164
DB 165 LGCEKNDVASYVPRPHNEMQIALIERGKCPFGDKSLAKGFETAVIYIDNPKSEGLH 244
OY 165 GAAOSDKGYSALVIGLSLEDGOKLI-KLAAGSVSYDLWVDSKQENRTYNNVAOTKGD 223
DB 245 GLGEPRTKHTVATVPYKVGKRLANILNDYSLYFAMDYVEFIKTQNTIADTKHGD 304
OY 224 PNNVVALGHTSVAGPGINDSGSIISNLVIAKALQYSKNAVRFLEWTAEEFGLLG 283
DB 305 PNIVIALGHSVSVEEGPINDSGSTISLLVNAKOLTFKNNKVRFWMAAEEBGLIG 364
OY 284 SNYVSHLATELNKIRLYLNEDMIASPNYALMIYDGDGSAFNSQAPGASQIEKLFEDY 343
DB 365 SNFYAYNLTKKEENSKIRVMDYDMASPNYEYID---ANNKENPKSEELKNLYVDY 420
OY 344 YDSIDLPHPITFDGRSDYEAFILNGIPSGGLFTGAEGIMSEENSRMGQGVAYDANY 403
DB 421 YKAHLNLTLPVEDRSDYVGFINGIPAGIATGAE---KNVNN---GKVLDRCY 471
OY 404 HAAGDMTNLNEAFLINSKATAFANATAYANDLSIPKENT 444
DB 472 HQLCDVSNLSWDATITYKLAHSVATYADSFEGPFRKET 512

RESULT 2
APX_STRGR STANDARD: PRT: 284 AA.
ID APX_STRGR STANDARD: PRT: 284 AA.
AC P80561:
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE AMINOPEPTIDASE (EC 3.4.11.-) (SGAP).
OS Streptomyces griseus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
RN (1)
RP SEQUENCE.
RX MEDLINE: 96270734.
RA Maras B., Greenblatt H.M., Shoham G., Spungin-Bialik A., Blumberg S.,
RA Barra D.,
RT "Amino-peptidase from Streptomyces griseus: primary structure and
RT comparison with other zinc-containing aminopeptidases";
RL Eur. J. Biochem. 236:843-846(1996).
RN (2)
RP SEQUENCE OF 1-6.
RX MEDLINE: 89338422.
RA Spungin A., Blumberg S.,
RT "Streptomyces griseus aminopeptidase is a calcium-activated zinc
RT metalloprotein. Purification and properties of the enzyme";
RL Eur. J. Biochem. 183:471-477(1989).

RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
RX MEDLINE: 97201142.
RA Greenblatt H.M., Almog O., Maras B., Spungin-Bialik A., Barra D.,
RA Blumberg S., Shoham G.,
RT "Streptomyces griseus aminopeptidase: X-ray crystallographic
RT structure at 1.75-A resolution";
RL J. Mol. Biol. 265:620-636(1997).

RN [4]
RP CHARACTERIZATION.
RX MEDLINE: 93185612.
RA Ben-Meir D., Spungin A., Ashkenazi R., Blumberg S.,
RT "Specificity of Streptomyces griseus aminopeptidase and modulation of
RT activity by divalent metal ion binding and substitution";
RL Eur. J. Biochem. 212:107-112(1993).
CC -1- FUNCTION: SPECIFIC FOR LARGER HYDROPHOBIC ACIDS, ESPECIALLY
CC LEUCINE. NO CLEAVAGE OCCURS IF THE NEXT RESIDUE IS PROLINE.
CC -1- COFACTOR: BINDS TWO ZINC IONS.
CC -1- ENZYME REGULATION: BINDS A CALCIUM ION WHICH MODULATES THE
CC ACTIVITY OF THE ENZYME.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- MASS SPECTROMETRY: MW=29728; MW_ERR=1.0; METHOD=ELECTROSPRAY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.
DR PDB: 1XJO: 01-APR-97.
KW Hydrolase; Aminopeptidase; Zymogen; Zinc; Calcium; 3D-structure.
FT METAL 85 85 ZINC 1.
FT METAL 97 97 ZINC 1 AND 2.
FT METAL 132 132 ZINC 2.
FT METAL 160 160 ZINC 1.
FT METAL 247 247 ZINC 2.
FT DISULFID 245 250
SO SEQUENCE 284 AA: 29722 MW: 671B80F8CA5C4CC CRC64:

Query Match 18.7%; Score 467.5; DB 1; Length 284;
Best Local Similarity 42.5%; Pred. No. 1,9e-23;
Matches 99; Conservative 40; Mismatches 79; Indels 15; Gaps 3;

OY 211 TTYNVVATKRGDPRNVVALGCHTDSVAGPGINDSGSIISNLVIAKALQ--YSKNA 268
DB 63 TGYNLIAWPGDGPKNYLMAGAHLDVSSGAGINDSGSAVETALVNSRAQYOPDKH 122
OY 269 VRFLEWTAEEFGLLGSNNYVSHLATELNKIRLYLNEDMIASPNYALMIYDGDGSAFNS 328
DB 123 LRFAMWGAEEGLIGSKRYVNNLPSADRSKLAGILNIDMISPGYFYDD----- 176
OY 329 GPAGSAQIEKLFEDYSDILP-HIPQFDGRSDYEAFILNGIPSGGLFTGAEGIMSEEN 387
DB 176 -----PVIEKTFKNVFAGLNPTLEIEEGDGRSDHAPFKNVGPVGGFTGAGYTKSAQ 230
OY 388 ASRNGQGVAYDANYHAAGDMTNLNEAFLINSKATAFANATAYANDLSIP 440
DB 231 AOKWGTAGQAFDRCHYSSCDLSINIDTALDRNSDAAHAIWTLSSGTGEP 283

RESULT 3
YVAD_BACSU STANDARD: PRT: 455 AA.
ID YVAD_BACSU STANDARD: PRT: 455 AA.
AC P25152:
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL 49.5 KD PROTEIN IN DAE-TYRZ INTERGENIC REGION PRECURSOR.
GN YVAD OR IPA-8R.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/staphylococcus group; Bacillus.
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN-168.
RX MEDLINE: 95020537.
RA Glaser P., Kunst F., Arnaud M., Coudart M.P., Gonzales W.,

RESULT

```
7
P2P_LACPA          P2E_LACPA      STANDARD:    PR1:   1902 AA.
ID                 ID                STANDARD:    PR1:   1902 AA.
AC              AC              AC              AC              AC              AC
DT       01-JUL-1993 (Rel. 26, Created)
DT       01-JUL-1993 (Rel. 26, Last sequence update)
DT       15-JUL-1998 (Rel. 36, Last annotation update)
DE     PIT-TYPE PROTEINASE PRECURSOR (EC 3.4.21.96) (LACTOCEPIN) (CELL WALL-ASSOCIATED SERINE PROTEINASE) (LP151).
DE     PRTP.
GN         Lactobacillus paracasei.
OS         Bacteria; Firmicutes; Bacillales/Clostridium group; Lactobacillaceae;
OC         Lactobacillus.
CC         [1]
RN         SEQUENCE FROM N.A.
RP         STRAIN-NCO 151;
RC         MEDLINE; 92381481.
RX         Holick A., Naes H.;
RT         "Cloning, sequencing and expression of the gene encoding the cell-envelope-associated proteolysin from Lactobacillus paracasei subsp. paracasei NCDO 151."
RT         J. Gen. Microbiol. 138:1353-1364(1992).
RN         [2]
RP         SEQUENCE OF 189-196.
RC         MEDLINE; 92226694.
RA         Naes H., Nissen-Meyer J.;
RT         "Purification and N-terminal amino acid sequence determination of the cell-wall-bound proteinase from Lactobacillus paracasei subsp. paracasei."
RL         J. Gen. Microbiol. 138:313-318(1992).
CC         -I- FUNCTION: PROTEASE WHICH BREAK DOWN MILK PROTEINS DURING THE GROWTH OF THE BACTERIA ON MILK.
CC         -I- CATALYTIC ACTIVITY: ENDOPEPTIDASE ACTIVITY WITH VERY BROAD SPECIFICITY, ALTHOUGH SOME SUBSITE PREFERENCE HAVE BEEN NOTED, E.G. LARGE HYDROPHOBIC RESIDUES IN THE P1 AND P4 POSITIONS, AND PRO IN THE P2 POSITION. BEST KNOWN FOR ITS ACTION ON CASEINS, ALTHOUGH IT HAS BEEN SHOWN TO HYDROLYSE HEMOGLOBIN AND OXIDIZED INSULIN B-CCHAIN.
CC         -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
CC         -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE SUBTILASAE FAMILY.
CC         -----
CC         THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation -
```

CC		the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (see http://www.isb.ch/announce/)	
CC	or send an email to license@isb.ch .	
CC	-----	
DR	EMBL; M83946; AAA25248.1; .	
DR	PIR; B44858; B44858.	
DR	HSSP; Q99405; IMPT.	
DR	PRINTS; PR00723; SUBTILISIN.	
DR	PROSITE; PS00136; SUBTILASE_ASP; 1.	
DR	PROSITE; PS00137; SUBTILASE_HIS; 1.	
DR	PROSITE; PS00138; SUBTILASE_SER; 1.	
DR	PROSITE; PS00343; GRAM_POS_ANCHORING; FALSE_NEG.	
DR	Pfam; PF00082; Peptidase_s8; 3.	
DR	Pfam; PF00746; Gram_pos_anchor; 1.	
KW	Hydrolase; Serine protease; Cell wall; Zymogen; Signal;	
KW	Transmembrane.	
FT	SIGNAL 1 33 POTENTIAL.	
FT	PROPEP 34 187 POTENTIAL.	
FT	CHAIN 188 1902 PII-TYPE PROTEINASE.	
FT	DOMAIN 188 1876 EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM 1877 1895 MEMBRANE ANCHOR (BY SIMILARITY).	
FT	DOMAIN 1896 1902 CYTOPLASMIC (POTENTIAL).	
FT	ACT_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).	
FT	ACT_SITE 281 281 CHARGE RELAY SYSTEM (BY SIMILARITY).	
FT	ACT_SITE 620 620 CHARGE RELAY SYSTEM (BY SIMILARITY).	
FT	DOMAIN 1867 1872 CONSERVED IN GRAM-POSITIVE COCCI SURFACE PROTEINS.	
SQ	SEQUENCE 1902 AA; 200253 MW; D8C9F36CE5DA582 CR664;	

Query Match	6.2%	Score 156;	DB 1;	Length 1902;
Best Local Similarity	21.3%	Pred. No. 0.027;		
Matches 109;	Conservative 60;	Mismatches 139;	Indels 204;	Gaps 26;

QY	80	LYKGDDEELEANTRYSPVE-----VTADVAVYKVLGSGSEADYPEDVKGKALLIRGCGP	134
Db	440	LQLGGETIQLSSNDPITGSSG-----FQDKKFTYVVKDASGDLSSKGA--ADYTADAKGKATAYKRGSLN	498
QY	135	FQDGSVLAAKAKAAASIYNNVAGSMAGTTLGAOSDKPYSAI-----VGISLEBQ	186
Db	499	FADKQKQYAGGAAGLLIVN-----DGTATPLTSIRLTTFPFPGSLSSKRG	546
QY	187	KLILIAEA--GSYSVDI-----WDSKQENRTTYNVAQTKGDPNNVALLGCH	234
Db	547	KLIVMYTAAHPDSDLGKIALTLPLPKQKYTEEDKMSDFTSYGVSNL--SFKPDITAPGNI	604
QY	235	DSVLEPGP-----	243
Db	605	WSTONNNNYTWSGTSMAPIIAGSALLKQALNNKNPFYADYKQLKGTALTDLKTVE	664
QY	243	-----IND-----DGSGLIENLYAKLQGYVKN-----AVRELF	273
Db	665	MNTAQPIINDIYNNVYSPRQAG---LVDYKKAIDALEKNSSTVVAENGYPRAVELKD	720
QY	274	WTAAE--FGLIGSNVYVSHL-----NATELKRILYLFNFDIMASENYALMTYD--DGA	324
Db	721	FTSIDKTEKFLFTNRTHTELYQMDSNINDYAN--YTSATDPNSG-VLYDKKIDGAA	774
QY	325	FNQSG---PAG-SAQIEKLEFDYYDSIDLPHIPTQDFGRSDYEAFI-----	367
Db	775	IKASGDIIVPAGKTAQIEFLS-----LPRKSFDOQFVEGELFNKSGDSGRNLTP	824
QY	367	-----LNGT---PSGGLFTAGAEIGMESENASRMGGQ--AGVADANTHA	405
Db	825	YMGFFGDWNDERKIVDSLIGITYTSPAGNY--GTVPLLTKNKT--GHQYGGWMTDA----	877
QY	406	AGDMNTNLNHEAFLNSKATAFAVATYANDIS	437
Db	877	--DGKQIVDDQALAFSSDKNAL-----YNDIS	901

	RESULT	8
CC	PSM_HUMAN	
AC	Q04609;	STANDARD; PRT; 750 AA.
DT	01-JUN-1994 (Rel. 29, Created)	
DT	01-JUN-1994 (Rel. 29, Last sequence update)	
DT	01-OCT-1996 (Rel. 34, Last annotation update)	
DE	PROSTATE-SPECIFIC MEMBRANE ANTIGEN (PSM).	
GN	PSM,	
OS	Homo sapiens (Human) .	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;	
OC	Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
RN	[1]	
RP	SEQUENCE FROM N.A.' AND PARTIAL SEQUENCE.	
RX	MEDLINE; 93113576.	
RA	Israel I.R.S., Powell C.T., Fair W.R., Heston W.D.;	
RT	"Molecular cloning of a complementary DNA encoding a prostate-specific membrane antigen."	
RL	Cancer Res. 53:227-230(1993).	
CC	-I- FUNCTION: MAY INTERACT WITH TRANSFERRIN OR ANOTHER LIGAND AND POSSIBLY FACILITATES METASTATIC SPREAD.	
CC	-I- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.	
CC	-I- TISSUE SPECIFICITY: PROSTATIC EPITHELIAL CELLS.	
CC	-I- SIMILARITY: TO THE TRANSFERRIN RECEPTOR PROTEIN.	
CC	CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its CC use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb.sib.ch).	

CC	EMBL:	M99487;	-	AAA60209.1;	-
DR	MTM:	600934;	-		
KY	Artisgen:	Transmembrane:	Glycoprotein:	Signal-anchor.	
FT	DOMAIN	1	19	CYTOSOLASMIC (POTENTIAL).	
FT	TRANSMEM	20	43	SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).	
FT	DOMAIN	44	750	EXTRACELLULAR (POTENTIAL).	
FT	MOD_RRS	1	1	BLOCKED.	
FT	CARBOHYD	76	76	POTENTIAL.	
FT	CARBOHYD	121	121	POTENTIAL.	
FT	CARBOHYD	140	140	POTENTIAL.	
FT	CARBOHYD	153	153	POTENTIAL.	
FT	CARBOHYD	195	195	POTENTIAL.	
FT	CARBOHYD	336	336	POTENTIAL.	
FT	CARBOHYD	459	459	POTENTIAL.	
FT	CARBOHYD	476	476	POTENTIAL.	
FT	CARBOHYD	638	638	POTENTIAL.	
FT	STILLAR	418	567	TO TRANSFEREIN RECEPTOR.	
FT	CONFLECT	750	354	R -> K (IN AA SEQUENCE).	
QQ	SEQUENCE	750 AA;	84330 MW;	ADBC0A7DBF47901A CRC64;	

Query Match Similarity	6.2%	Score 154	DB 1	Length 750
Best Local Similarity	20.1%	Pred. No. 0.011		
Matches 90; Conservative	59;	Mismatches 170;	Indels 188;	Gaps 17;

QY	28	DFAAYAYERNRNVFCGRADDTVNYLYE---	ELKKTGYDYVKQOQVHLMSNADQTLKVG	84
Db	114	DVLISYD-----NKHHPNISTINEDGNEIENTSLF---	EBPPGCEYENV-----	155
QY	85	EEIEAKTMTSPSYEYVADVAVY-----	KNLGCSEADYPSDVEKVALIKRGCECPFGDKS	139
Db	155	SDIYPPPSAASPOGMEPGDLVYVNTATEDFEFLERMDKINCCKITVIRKGYFRGNKV		214
QY	140	VLAAKAKAASISYVNNVA-----	GSNAGTL-----	165
Db	215	KNMOLAGAKGVILYSDPADYFAPGVKSYPDGMWLLPGGGVORGNTLNLGAGDPLTPGPYA		274
QY	165	-----GAQSDKGPYSALVIGTISLEBDGKLIK-----		191
Db	275	NEYAYRRGILAEVAGLPSIPVHPHIGIYDAOKLLLEKMGSSAPDSSWRGSLKVPYVNVGFGTT		334

```

QY 192 AAGSVSDVLDWDSKQENTTYNVAQTKGS-DPNNVVALGGHTDSVEAGPSIN-DDSGS 249
DQ 335 GNFSTQKVMHSHSTNETRITNVLGAVEPDRYVLGGHRSWFG-GIDPSGAA 393
DQ 250 IISNIVAKALTO---YSVKNVREFLEWTAEEFGLGSNYVYSHLNATELNKIRLYLND 306
DQ 394 VVHEIVRSFGTLAKKGGWRRTILFASWDAEEFGLGSTEMAENSRLQERGVAYIND 453
QY 307 MIASPNVAMLYDGDGSAFNOS-----GPAGSAQIEKLFEDYDIDLPHPHQF 356
DQ 454 SEIEGNTYLRV---DCTPLMTSLVHNLTKELKSPDEGFGKSLYSWTKKSPSP-----EF 506
QY 357 DGR-----SDYEAFILN-GIPSG 373
DQ 507 SGMPRIKSLKSGNDVEFFORLGIASG 533

RESULT 9
ID YBS4_YEAST STANDARD: PRT: 413 AA.
AC P36244;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOTHEICAL 47.8 KD PROTEIN IN HSP26-TIF32 INTERGENIC REGION.
GN YBR074W OR YBR0718.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C;
RX MEDLINE; 95076715.
RA van der Aart Q.J.M., Barthe C., Daignon F., Aigle M., Cruzet M.,
RA Steensma H.Y.;
RT "Sequence analysis of a 31 kb DNA fragment from the right arm of
RT Saccharomyces cerevisiae chromosome II.";
RL Yeast 10:959-964(1994).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M33; ALSO KNOWN AS THE
CC AMINOPEPTIDASE Y FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X76394; CAAS3931.1; -.
DR EMBL; Z35943; CAAS5018.1; -.
DR PIR; S45467; S45467.
DR PIR; S39224; S39224.
KW Hypothetical protein; Hydrolase; Aminopeptidase; Zinc;
KW Metalloprotease.
SQ SEQUENCE 413 AA: 47802 MW: 469B359A66AB621C CRC64:

```

```

DQ 235 VLFRSDSTARLYQSVKKNFNGSIYQGFYSRY-----VRSEDTYKYEEN 283
QY 369GI-----PSGLFTGAGCINSEENASRW 391
DQ 284 GMRGMDVAFYKPRNLVHTIKDSIQYTSKASLW 315

RESULT 10
ID PIP_LACLC STANDARD: PRT: 1902 AA.
AC P16271;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE P1-TYPE PROTEINASE PRECURSOR (EC 3.4.21.-) (WALL-ASSOCIATED SERINE
DE PROTEINASE).
GN PRTP.
OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OG Plasmid pMW05.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WG2;
RX MEDLINE; 88149035.
RA KOK J., Leenhouts K.J., Haandrikman A.J., Ledebor A.M., Venema G.;
RT "Nucleotide sequence of the cell wall proteinase gene of
RT Streptococcus cremoris Wg2.";
RL Appl. Environ. Microbiol. 54:231-238(1988).
CC -1- FUNCTION: PROTEASE WHICH BREAK DOWN MILK PROTEINS DURING THE
CC GROWTH OF THE BACTERIA ON MILK.
CC -1- CATALYTIC ACTIVITY: ENDOPEPTIDASE ACTIVITY WITH VERY BROAD
CC SPECIFICITY, ALTHOUGH SOME SUBSITE PREFERENCE HAVE BEEN NOTED,
CC E.G. LARGE HYDROPHOBIC RESIDUES IN THE P1 AND P4 POSITIONS, AND
CC PRO IN THE P2 POSITION. BEST KNOWN FOR ITS ACTION ON CASEINS,
CC ALTHOUGH IT HAS BEEN SHOWN TO HYDROLYSE HEMOGLOBIN AND OXIDIZED
CC INSULIN B-CHAIN.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M24767; AAA17677.1; -.
DR HSRP; 099405; IMPT.
DR PRINTS; PRO0723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; 1.
DR PIRAM; PF00082; Peptidase_S8; 3.
DR PIRAM; PF00746; Gram_pos_anchor; 1.
KW Hydrolase; Serine protease; Cell wall; Zymogen; signal; Plasmid;
KW Transmembrane.
FT SIGNAL 1 33
FT PROPEP 34 187
FT CHAIN 188 1902
FT DOMAIN 188 1876
FT TRANSMEM 1877 1895
FT DOMAIN 1896 1902
FT ACT_SITE 217 217
FT ACT_SITE 281 281
FT ACT_SITE 620 620
FT DOMAIN 1867 1872
FT SEQUENCE 1902 AA: 199910 MW: 2901C7F19B2E5D0B CRC64:

```


DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE CSA PEPTIDASE PRECURSOR (EC 3.4.21.-) (SCP).
 GN SCPA.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 CC Streptococcus.
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-41.
 RX MEDLINE; 90153964.
 RA Chen C.C., Cleary P.P.;
 RT "Complete nucleotide sequence of the streptococcal Csa peptidase gene
 RT of Streptococcus pyogenes.";
 RL J. Biol. Chem. 265:3161-3167(1990).
 CC -!- FUNCTION: THIS VIRULENCE FACTOR OF S.PYOGENES SPECIFICALLY CLEAVES
 CC THE HUMAN SERUM CHEMOTAXIN CSA AT LYS(68)-ASP(69) BOND NEAR ITS
 CC C-TERMINUS, DESTROYING ITS ABILITY TO SERVE AS A CHEMOTACTANT.
 CC -!- SUBCELLULAR LOCATION: CELL WALL BOUND.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
 CC SUBTILASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; J05229; AAA26960.1; -.
 DR PIR; A35066; A35066.
 DR HSSP; P00782; 2SBT.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_LASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING; FALSE_NEG.
 DR PRAM; PF00082; Peptidase_S8; 4.
 DR PRAM; PF00746; Gram_pos_anchor; 1.
 KW Hydroxylase; Serine protease; Signal; Cell wall; Repeat; Transmembrane.
 FT SIGNAL 1 31
 FT CHAIN 32 1167
 FT DOMAIN 32 1139
 FT TRANSMEM 1140 1157
 FT DOMAIN 1158 1167
 FT ACT_SITE 130 130
 FT ACT_SITE 193 193
 FT ACT_SITE 512 512
 FT DOMAIN 1029 1104
 FT DOMAIN 1034 1101
 FT REPEAT 1034 1050
 FT REPEAT 1051 1067
 FT REPEAT 1068 1084
 FT REPEAT 1085 1101
 FT DOMAIN 1127 1131
 FT
 FT SEQUENCE 1167 AA; 128263 MW; D2DDCS2E5752DA5D CR64;
 SO
 Query Match 6.0%; Score 149; DB 1; Length 1167;
 Best Local Similarity 21.9%; Pred. No. 0.04;
 Matches 121; Conservative 66; Mismatches 196; Indels 170; Gaps 26;
 QY 6 SPDEFPPDIOLEDLBESQ-----QLEDFAYIYPER-----NRVGGK 43
 DB 205 APSEKPEPRLEGAMPKAOQLIMREIVNGIADYARNYAQIRDAVNLGAKVIMSGNA 264
 QY 44 AHDQTVNY--LYEELKTKGYDYVKQPOVHLSMNAQDTLK-----VGD 84
 DB 265 A-----LAVANLPDETAKA--FDYAKSKGVSLVTSAGNDSFGGKTRPLPADHPDYGVGT 318
 QY 85 EELEAKMT--YSPSVEVTADVAVVR-----NLGCS 113

DB 319 PAADSTILTVASYSRQOLF-ETAMVKTDDQOKEMFVLSINREPRKAYDAVANGMK 377
 QY 114 EADPSVEKGVALLIKGECPFGDKSVLAKAKAAASIVNNAGSAGLGAQSGK- 173
 DB 378 EDDF-KDVKRGIALIERGIDIDFKDVANKAKAGAVGLIYDN-----ODKG 423
 QY 173 -----PYSAIVGISLEDDGOKLIKLAAGSVSVDLVWDSKOEKRTYNY--VAOTKGG 222
 DB 424 PLEPNDQMPAATISKDG-----LLKPNPQTTTFNATPVLPYASG 468
 QY 223 DPNVVALGHTDSVEAPGINDGSGIISNLVIK--ALTOYSKNAVLEFLWTAEEFG 280
 DB 469 TKLSRFSSWGLTAGCNLIKPDIAAGQDILLSVANAKKAKLSGMSAPL-----VAGIMG 523
 QY 281 LLSGNYVSHLNTLKLIRLYLNFDMIASPNALMYDGDGSAF--NOSGPGASQITE 337
 DB 524 LLOKQYETQPDMPSPRLDLAKKVLN-----SSATALYDEDEKAYFSPROOG-AGAVDAK 578
 QY 338 KLF-----EDYDSIDLPHIPIQFD-----GRSD-----YEAFLINGIPSGGLF 376
 DB 579 KASATATVTDKNTSKSVHLNVSDFEYTVVHNKSDKPPQELYQATVQTDKVDGKH 638
 QY 377 TGAEIGNSE-----ENASRWGQAGAYADANYHAAGDNMTNLNHEAFLINSKATAF 427
 DB 639 ALAKVLYEASWOKITIPANSK---QVTVPIDAS-RFSNDLLAQMKNGYFL--EGFVRF 692
 QY 428 AVATYANDLSIP 440
 DB 693 KQDPTKEELMSIP 705
 RESULT 13
 Y96_YEAST
 ID Y96_YEAST STANDARD; PRT; 811 AA.
 AC P47161;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE HYPONETICAL 92.0 KD PROTEIN IN RPSS-2MS1 INTERGENIC REGION.
 GN YR126C OR J2050.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 CC Saccharomycetaceae; Saccharomycetes.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rose M., Koeltter P., Entian K.D.;
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: TO MAMMALIAN TRANSFERIN RECEPTOR PROTEIN AND
 CC PROSTATE-SPECIFIC MEMBRANE ANTIGEN (PSM).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; Z49626; CA89657.1; -.
 KW Hypothetical protein.
 SO SEQUENCE 811 AA; 92017 MW; 932818225880E49B CR64;
 QY
 Query Match 5.3%; Score 133.5; DB 1; Length 811;
 Best Local Similarity 18.9%; Pred. No. 0.24;
 Matches 111; Conservative 76; Mismatches 207; Indels 193; Gaps 25;
 QY 20 LEGSQOLEDAVAYVPERNRYFGKKAHDYVNYLYEELKTKGYDYVKQPOV---HLW--- 74
 DB 139 LQENRAKEHYVTK--AGVSNQASDSSTFKYLLDELDMGY-----KPEVEYIYPIGE 191
 QY 74 -----SNADQTLKGVDEIEAKTM-----TSPSVEVTADVAVVKNL 110

DB 192 PYDTNAPLENGKVVYEASMIEDRVK-GDPASHARKRGKGFHOYSKNGSVTARY-VFCNY 249
 QY 111 GCSEADYP-----SDVEKVALIKRGCEPFGRDYSVLAKAKAASIYN----- 155
 DB 250 G-SISDYKLLKLNKIDIEDKIHIVRSKILPGIKVNAELYGASVYIITDFEDDKYTE 308
 QY 155 -----NVAGSMAG--TIGAAQSD-----KGPYSATV 178
 DB 309 ENGFLHYPGAPANPSYIRRDVYNFSDTPGDPPTYGYSKSDTEHMSPGVGPPIPSV 368
 QY 179 GISLEDGQKLIK-----LABAGSY-----SVDLVWSDKOENRTTYNVAQT 219
 DB 369 PMSARARVQPIELERNRGFOIGRSGIKRPGSFSTGSSSID---KYHLMNELTYNKKMS 425
 QY 220 K-----GGDPNNVVALGCHTDSVEAGPGINDGSGIISNLYAKALQ-----YSKVA 268
 DB 426 SVEVSIPIGTFEGEIIIGAHKDSL-ASSAGDANSQSAILLETARQMSKILKHGMPLRP 484
 QY 269 VFLEFMTAEFGILGSGNYSVSHLNATELUNKIRLYLFEDIASPNVYLMYDDGSAFNOS 328
 DB 485 IKLISMDGSRSGLGSTDAEAHAALRRRALVYLLDNVAIS-----GTFHF-- 532
 QY 329 GPAGSNOIEKLEFD-YVDSIDLPHPITOPDGRSDYEA-----ILNGIPSGGL 375
 DB 532 -----CKANPLLDQVYIEAKL---TEFGHEDMSLFPHWKYSNATISLDDGLSYTS 582
 QY 376 FPGAGCIMSSEENASRWGQAGVAYDAN-----YH 404
 DB 583 FOYHLGVPAAHFOFNANDYSAGVYHSNSVFDSPTWLEKFTNSDYKLHNTAMFVGLTILM 642
 QY 405 AAGDMNTNLNHEAFILNSKATAFAVATYANDLSIPKRTTSLHRR 451
 DB 643 LSENELARNTHYLL-KKIYMYIAMSNSLSAFAFDDEVNSLAKR 687

RESULT 14
 TRSR_HUMAN
 ID TRSR_HUMAN STANDARD: PRT: 760 AA.
 AC P02786:
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TRANSFERRIN RECEPTOR PROTEIN (TR) (ANTIGEN CD71) (T9) (P90).
 GN TFR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 85012743.
 RA Schneider C., Owen M.J., Banville D., Williams J.G.;
 RT "Primary structure of human transferrin receptor deduced from the
 RT mRNA sequence.";
 RL Nature 311:675-678(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 85048936.
 RA McCelland A., Kunh L.C., Ruddle F.H.;
 RT "The human transferrin receptor gene: genomic organization, and the
 RT complete primary structure of the receptor deduced from a cDNA
 RT sequence.";
 RL Cell 39:267-274(1984).
 RN [3]
 RP SEQUENCE OF 101-119 (SERUM FORM).
 RX MEDLINE: 9103436.
 RA Shih Y.J., Bayns R.D., Hudson B.G., Flowers C.H., Skikne B.S.,
 RA Cook J.D.;
 RT "Serum transferrin receptor is a truncated form of tissue receptor.";
 RL J. Biol. Chem. 265:19077-19081(1990).
 RN [4]
 RP PALMITOYLATION OF CYS-62.
 RX MEDLINE: 87218484.
 RA Jting S., Trowbridge I.S.;

RT "Identification of the intermolecular disulfide bonds of the human
 RT transferrin receptor and its lipid-attachment site.";
 RL EMBO J. 6:327-331(1987).
 RN [5]
 RP MUTAGENESIS OF CYSTEINES INVOLVED IN INTERMOLECULAR BONDS.
 RX MEDLINE: 90005427.
 RA Alvarez E., Girones N., Davis R.J.;
 RT "Intermolecular disulfide bonds are not required for the expression
 RT of the dimeric state and functional activity of the transferrin
 RT receptor.";
 RL EMBO J. 8:2231-2240(1989).
 RN [6]
 RP INTERNALIZATION SEQUENCE, AND MUTAGENESIS OF TYR-20.
 RX MEDLINE: 90130619.
 RA Jting S., Spencer T., Miller K., Hopkins C., Trowbridge I.S.;
 RT "Role of the human transferrin receptor cytoplasmic domain in
 RT endocytosis: localization of a specific signal sequence for
 RT internalization.";
 RL J. Cell Biol. 110:283-294(1990).
 RN [7]
 RP MUTAGENESIS.
 RX MEDLINE: 94012749.
 RA Collawn J.F., Lai A., Domingo D., Fitch M., Hatton S.,
 RA Trowbridge I.S.;
 RT "YTRF is the conserved internalization signal of the transferrin
 RT receptor, and a second YTRF signal at position 31-34 enhances
 RT endocytosis.";
 RL J. Biol. Chem. 268:21686-21692(1993).
 RN [8]
 RP CARBOHYDRATE-BINDING SITES THR-104.
 RX MEDLINE: 93043836.
 RA Do S.I., Cummings R.D.;
 RT "Presence of O-linked oligosaccharide on a threonine residue in the
 RT human transferrin receptor.";
 RL Glycobiology 2:345-353(1992).
 RN [9]
 RP CARBOHYDRATE-BINDING SITES THR-104.
 RX MEDLINE: 93043837.
 RA Hayes G.R., Enns C.A., Lucas J.J.;
 RT "Identification of the O-linked glycosylation site of the human
 RT transferrin receptor.";
 RL Glycobiology 2:355-359(1992).
 RN [10]
 RP STRUCTURE OF CARBOHYDRATE ASN-727.
 RX MEDLINE: 95289226.
 RA Hayes G.R., Williams A., Costello C.E., Enns C.A., Lucas J.J.;
 RT "The critical glycosylation site of human transferrin receptor
 RT contains a high-mannose oligosaccharide.";
 RL Glycobiology 5:227-232(1995).
 CC -1- FUNCTION: CELLULAR UPTAKE OF IRON OCCURS VIA RECEPTOR-MEDIATED
 CC ENDOCYTOSIS OF LIGAND-OCCUPIED TRANSFERRIN RECEPTOR.
 CC -1- SUBUNIT: DIMER OF SIMILAR OR IDENTICAL CHAINS LINKED BY TWO
 CC DISULFIDE BONDS. A MONOMERIC SERUM FORM OF THE RECEPTOR EXISTS, IT
 CC LACKS RESIDUES 1-100.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.
 CC -1- PTM: EACH CHAIN BINDS GLYCANS, PHOSPHATE, AND A PALMITATE GROUP.
 CC -1- SIMILARITY: TO PROSTATE-SPECIFIC MEMBRANE ANTIGEN (PSM).
 CC -1- DATABASE: NAME-PROT; NOTE-CD guide CD71 entry?
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd71.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X01060; CAA25527.1; -;
 CC EMBL: M11507; AAA61153.1; -;
 CC DR PIR: A03259; JXHU.
 CC DR MIM: 190010; -;
 KW Transmembrane; Glycoprotein; Receptor; Lipoprotein; Palmitate;

Signal-anchor; Endocytosis.
CYTOPLASMIC (POTENTIAL).
SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
EXTRACELLULAR; CONTAINS TRANSERRIN
BINDING SITE (POTENTIAL).
STOP-TRANSFER SEQUENCE.
PALMITATE.
INTERCHAIN.
INTERCHAIN.
PSNL->YTRF: ONLY 80 % AS ACTIVE AS
WILD-TYPE RECEPTOR.
Y->C: ONLY 35% AS ACTIVE AS WILD-TYPE
RECEPTOR.
Y->G: ONLY 20% AS ACTIVE AS WILD-TYPE
RECEPTOR.
YTRF->PEPG: ONLY 16% AS ACTIVE AS
WILD-TYPE RECEPTOR.
YTRFSLAROVGDMS->PPEGLSLAROVYTRF: NO
INFLUENCE ON ENDOCYTIC UPTAKE OF THE
RECEPTOR.
YTRF->YTRF: ONLY 14% AS ACTIVE AS
WILD-TYPE RECEPTOR.
YTRF->YTRF: ONLY 19% AS ACTIVE AS
WILD-TYPE RECEPTOR.
T->F: ONLY 88% AS ACTIVE AS WILD-TYPE
RECEPTOR.
F->Y: ONLY 48% AS ACTIVE AS WILD-TYPE
RECEPTOR.
GDMS->YTRF: 2-FOLD INCREASE OF THE
ENDOCYTIC UPTAKE OF THE RECEPTOR.
NADN->YTRF: 1.27-FOLD INCREASE OF THE
ENDOCYTIC UPTAKE OF THE RECEPTOR.
ENDOCYTIC SIGNAL.
SEQUENCE 760 AA; 84901 MW; 4AAE97675BA9F2D CRC64;

Query Match 5.2%; Score 130; DB 1; Length 760;

Best Local Similarity 20.1%; Pred. No. 0.38; Mismatches 130; Indels 98; Gaps 15;

Matches 72; Conservative 58; Mismatches 130; Indels 98; Gaps 15;

QY 92 MTYPSVEVTADVAVNKNGCSE--ADYPSVEGVKVALIKRGECPGDKSVYAAKAKAAA 149
D 220 VAYSKATVITGKL-VHANFGTKKDFEDLYTPVNGSIVYRAGKITFAEKVANAESLNAIG 278
QY 150 SIYNN-----VAGSMGTGAOSDKG-PYS-----AIVGIS 181
D 279 VLIYMDQTKPTVNAHELSFEGHAGLGTDPYTPGPPSNHTQFPSSSGGLPNIPVQGIS 338
QY 182 LEDQOKLI-----KLAAGSVSYDLTAVDSKOENRTYVNAOTKG-CD 223
D 339 RAAAEKLFEGNNEGPCSDMTKSTCRMTSESKNKVLVSVALKEIKLINFVIGKVE 398
QY 224 PNNVVALGHTDSVEAGCGINDDSGIIISNVIKALTOYSVKN-----AVRELEWTA 277
D 399 PDHVVVVAQSDA--WCGGAAGSGVGTALLKLAOMFSDMLKDGQFQPSRSIIIFASWSAG 456
QY 278 EFGLLGSNYYVS-----HLNA-TELNKRIRLYL--NFDMLASPYALMIYKDGSAANO 327
D 457 DFGSVGATEWLEGLSLSHLKAFTYINLDKAVLIGTSNKVVASPLLYLL--EXTMONV 513
QY 328 SGPA-----GSAQIEKLFED-----YYSDIDLPHIPTOFD 357
D 514 KHPYTGGLYODSNMWSAVEKLTLDNAFPFLAYSGIPAVSFCEODDIDYRILGTMD 571

RESULT 15
VPH_BPHPI STANDARD; PRT; 925 AA.

AC P51735;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE PROBABLE TAIL FIBER PROTEIN (ORF31).
OS Bacteriophage HPI.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HPIC1;
RA MEDLINE; 96279738.
RA Esposito D., Fitzmaurice W.P., Benjamin R.C., Goodman S.D.,
RA Waldman A.S., Scooca J.J.;
RL "The complete nucleotide sequence of bacteriophage HPI DNA.";
RL Nucleic Acids Res. 24:2360-2368(1996).
CC -1- SIMILARITY: SOME, TO PHAGE P22 PROTEIN H.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U24159; AAB09218.1; -.
KW Fiber protein.
SQ SEQUENCE 925 AA; 102001 MW; FB2C99ED9FDB090E CRC64;

Query Match 5.2%; Score 129; DB 1; Length 925;

Best Local Similarity 21.2%; Pred. No. 0.57; Mismatches 226; Indels 174; Gaps 27;

Matches 126; Conservative 66; Mismatches 226; Indels 174; Gaps 27;

QY 1 GRALVSDPEDEPDIQDLDELGSQLEDFAVYPERNRVYEGKADDTVNL----- 53
D 169 GRAVFFDQSF-----LVKRTKGNQFTIQPGNAVYEGVRMDLGTGHTLHANSIPESIVADY 223
QY 53 -----YELKKTGYD-VYKQPOVHMSNADQTLKVDDELEA-----KT 91
D 224 VHCITGEGYQTEIKYLTQSKADYVDYANRQVYQIADIDSQGNVYDRRLSPFGMNP 283
QY 92 MTYPSVEVTAD-----VAVVNKNGCSEADYP--SPVEGVKVALIKRGECPGDKSV 140
D 284 LTLDITENTNDKGRHKLPISIKRGTIKLSANSDSETOAASK----- 333
QY 141 LAARAKAASIVYNNVAGSMAGTLG--AAOSDKQYSAIV--GISLEDQOKLIK--LA 192
D 333 -AVKTAADKAVEYKTAESKVGKLGNGNESIQGTKEFKIIGFRIGVADSGTYANAHLL 391
QY 193 EAGSVSYDLTAVDSKOENRTYVNAOTKG-----GDPNNVA 229
D 392 NMGANDDQWIEYKRSKRVGTIRIRANGELSYNNOKIYHAGAKPONTDIEGRPTIAG 451
QY 230 LGHTDSVEAGP-----GINDGSGIIS--NL-----VIKALTOYSKVAPELF 273
D 452 YGIGNFYEDQOGANGYKTDGNTYTLASGQNLPENGEHLIEVSGGAT-----NAVQIA 506
QY 274 WTAEFGLL-----GSNYYVSHLNATELNKRIRLYLND-MIASPNVYALMIYDGGSAFNQ 327
D 507 RKANDNKIKTRFFNGSNWSEKWDAGDGVPLGVSVPRAVTN---VGLKANGTFFNQ 563
QY 328 SGPAAGQIEKLFEDY-----DSIDLPHIPOPGRBDDYEFILNGIPSGGL-FTGAEI 382
D 564 -----QTFPDLIRLIGDSSQDLPLRSDVGMATY--FAVDNIPSGVIADFSIRST 611
QY 383 MSEENASRWGQAGVADANYHAAGDMNTNLN-----EAFILNS-----K 423
D 612 VTQON-----YPELYQYLVDKYSSISNPLADEDFIRNTGNGNLNGTQSDPIK 660
QY 424 ATAAVAVATYANDL--SSIPKRNITSLHRAARTMRPFGKRAPKTHAVSGSGCNH 476
D 661 KHVARVATHWADSSDSIFYDKTKTVIDSRLT-----ATTDDMLSDNGFHH 708

Thu May 11 10:03:19 2000

us-09-080-127-2_copy_16_496.rsp

Page 12

Search completed: May 10, 2000, 20:50:15
Job time: 4623 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2000, 19:19:47 ; Search time 50.18 seconds

(without alignments)
664.602 Million cell updates/sec

Title: US-09-080-127-2_COPY_16_496

Perfect score: 2500
Sequence: 1 GRALVSPDFEDIDLEDDL.....APKTHAVSGSACHSQVEA 481

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SPREMBL_12:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	668	26.7	500	2	P96264 mycobacteri
2	527.5	21.1	485	2	Q53737 streptomyc
3	229.5	9.2	472	4	Q9V646 homo sapien
4	220.5	8.8	472	11	Q9Z1Y1 rattus norv
5	219.5	8.8	541	4	Q9Y5X6 homo sapien
6	215.5	8.6	470	11	Q9WVJ3 mus muscul
7	207.5	8.3	493	5	Q76552 acanthochei
8	189	7.6	368	2	Q67732 aquifex aeo
9	166	6.6	752	11	P70627 rattus norv
10	164.5	6.6	433	11	P70216 mus muscul
11	163	6.5	751	5	P91406 caenorhabd1
12	158	6.3	565	1	Q57813 pyrococcus
13	158	6.3	751	6	Q77564 sus scrofa
14	157.5	6.3	393	2	Q82996 aeromonas c
15	157.5	6.3	374	3	Q04033 saccharomyc
16	154.5	6.2	1483	5	Q93332 caenorhabd1
17	154	6.2	750	4	Q43748 homo sapien
18	149.5	6.0	757	11	Q07891 cricetus
19	149	6.0	763	11	Q62351 mus muscul
20	149	6.0	778	13	Q90997 gallus gall

21	148.5	5.9	752	11	Q35409 mus muscul
22	148	5.9	1150	2	Q53637 streptococ
23	140.5	5.6	622	11	Q99376 rattus norv
24	136	5.5	740	4	Q9Y3Q0 homo sapien
25	136	5.4	794	3	Q43023 schizosach
26	132.5	5.3	745	11	Q54697 rattus norv
27	130.5	5.2	416	2	Q9X6X2 myxococcus
28	130	5.2	501	2	P96152 vibrio chol
29	129.5	5.2	780	2	Q9ZEK2 anabaena va
30	125	5.0	782	4	Q75422 homo sapien
31	117	4.7	1035	13	Q57537 xenopus lae
32	115.5	4.6	467	2	Q84914 lactobacill
33	114.5	4.6	1770	2	Q84419 chlamydia t
34	114	4.6	2614	5	Q97054 dictyostell
35	112	4.5	1685	10	Q04142 porphyra ye
36	109.5	4.4	567	1	Q9YDB0 aeropyrum p
37	109	4.4	505	2	Q57381 salmonella
38	109	4.4	1148	2	Q60045 thermoaer
39	108	4.3	428	2	Q06803 mycobacteri
40	108	4.3	494	2	Q54210 salmonella
41	107.5	4.3	852	2	P76087 escherichia
42	107.5	4.3	852	2	P76856 escherichia
43	107	4.3	505	2	Q54863 salmonella
44	107	4.3	505	2	Q54864 salmonella
45	107	4.3	690	1	Q26901 methanobact

ALIGNMENTS

RESULT 1
ID P96264 PRELIMINARY; PRT; 500 AA.
AC P96264;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DI 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE HYPOTHETICAL 52.0 KD PROTEIN.
GN MTCY22G10.15.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RA SKELTON J., CHURCHER C.M.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RA BARRELL B.G., RAJANDREAN M.A.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE: 96181548.
RX PHILLIPS W.J., POULET S., EIGMEIER K., PASCOPELLA L.,
RA BALASUBRAMANIAN V., HEYM B., BERGER S., BLOOM B.R., JACOBS W.R. JR.,
RA COLE S.T.;
RT "An integrated map of the genome of the tubercle bacillus,
RT Mycobacterium tuberculosis H37RV, and comparison with Mycobacterium
RT lepre.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:3132-3137(1996).
DR EMBL: Z84724; CAB06558.1; -.
DR HSSP: P80561; IXJO.
KW Hypothetical protein.
SQ SEQUENCE 500 AA; 52042 MW; 9AC2ADA2 CRC32;

Query Match 26.7%; Score 668; DB 2; Length 500;
Best Local Similarity 38.5%; Pred. No. 3, 3e-38;
Matches 170; Conservative 57; Mismatches 184; Indels 30; Gaps 11;

```

OY 14 IQLEDLLESGSOOLEDEFAVAPERNVFGSKAHDDTVNLYELKTKGYDYVKOPQVHLM 73
DB 52 VSTDMMHLSLQDIANA-NDGTRAVGTPG;QASVDIVYNTLRNSGEVOTPEESARFV 110
OY 74 SNADOTLKVGDEIEBAKTUTYS-----PSVEVTAADVAVNNLGCSEADYPS-DVE 122
DB 111 KAEKGVVTLGGTVTERALLEYSLGTPPDGVTPVLAAPAD-----DSPGSPSDYRLPVS 166
OY 123 GKVALIKRGCEPFGDKSVLAAKAKAASTIVNNV-AGSMAGTLGAASQDKPYSAIVGIS 181
DB 167 GAVVIVYDVRGVCFAQKEDAAORGAVALIADNIDEOAAGGTILGANTDKIP---VVSAT 223
OY 182 LEDGOKLKLIAAGSVYDVLMTWDSKOENRTTYNVAOTGDDPNNVVALGHTDSVEAGP 241
DB 224 KSVGQOL--RQSGSPTTKYL--TASTQSFKARNVTAQTGSSANVYVMAAHLDSVEEP 279
OY 242 GINDSGSIISNLVIKAL-TOYSVKNAYRFLFWTAEEFGLGSNYYVSHLNATELNKIR 300
DB 280 GINDSGSVAALVETAVOLGNSPHVSNAYRFAFWGAEFGLGSNYYVESLDIDALKGIA 339
OY 301 LYLNMDIASPNVYALMTIDGDS-AFNOSG----PAGSNOIKLEEDYDSDIDLPHITPQ 355
DB 340 LYLNMDIASPNVYALMTIDGDS-AFNOSG----PAGSNOIKLEEDYDSDIDLPHITPQ 399
OY 356 FDGRSDYEAFILNGIPSGGLFTGAGIMSEENASRMGQAGVAYDANYHAADNNTNLNH 415
DB 400 FDGRSDYOGFTLAGIPSGGLFTGAGIMSEENASRMGQAGVAYDANYHAADNNTNLNH 459
OY 416 EAFILNSKATAPAVATYANDL 436
DB 460 TALGINGAGVAVAGLYAODL 480

RESULT 2
OY 053737 PRELIMINARY: PRT: 485 AA.
AC 053737;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE N-ACETYLPUROMICIN N-ACETYLHYDROLASE PRECURSOR.
GN NAPH.
OS Streptomyces lipmanii (Streptomyces alboniger).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC12461.
RX MEDLINE: 96139493.
RA TERCERO J.A., ESPINOSA J.C., LACALLE R.A., JIMENEZ A.;
RT "The biosynthetic pathway of the aminocyclitol antibiotic puromycin,
as deduced from the molecular analysis of the pur cluster of
Streptomyces alboniger."
RL J. Biol. Chem. 271:1579-1590 (1996).
DR HSPG; X92429; CA63158.1; -.
KW Signal: Hydrolase.
FT SIGNAL 1 33 POTENTIAL.
FT CHAIN 34 485
SO SEQUENCE 485 AA; 51358 MM; E2BCE65 CRC32;
```

Query Match 21.1%; Score 527.5; DB 2; Length 485;
Best Local Similarity 33.6%; Pred. No. 1.4e-28;
Matches 151; Conservative 67; Mismatches 206; Indels 25; Gaps 11;

```

OY 1 GRALSPDEFEDIQLEDLESGSOOLEDEFAVAPERNVFGSKAHDDTVNLYELKTKTG 60
DB 34 GSEVETPPPLSOD-RIEKHLTFERLAD-AHC---GDRASGSGAGAESVDYIEKLIKAG 88
OY 61 YDYVYKOPQVHLSNADOTLKVGDEIEBAKTUT-YSPSVE---VTADAAVY-----KILG 111
DB 89 YSTGQOCPFPFRYTRLEKLVLRDSTPGVVVSGYSTPGRGLTARPTAVAGDTERROG 148
```

```

OY 112 CSEADYPS-DVEGKVALIKRGCEPFGDKSVLAAKAKAASTIVNNVNASMAGTLGAASD 170
DB 149 CRAGAYDGVPAAGRIALVADGAGCSNDDKORVAAGAALVAVANTGPELTTWLADPEAA 208
OY 171 KGPYSIVGISELEDGOKLIKLAAGSVYDVLMTWDSKOENRTTYNVAOTKGGDPNNVAL 230
DB 209 RLP---IGVYQERGRLLAARAGR-TVOJTLISLTERRTTEMLIANSPRNGEHHYVS 264
OY 221 GGHDSVEAGCINDGSG----IISNLVIKALTOYSVKNAYRFLFWTAEEFGLGSNTY 286
DB 265 GAHLDSVEPEGGINNGVAAVLLLESALAAKDHGTRAEEDRLVFGFWGAEFGLGSQH 324
OY 287 YVSHLNATELNKIRLYNFMEDIASPNVYALMTIDGDS--AFNOSGPGASQIEKTFEDY 344
DB 325 YVSHLNATELNKIRLYNFMEDIASPNVYALMTIDGDS--AFNOSGPGASQIEKTFEDY 384
OY 345 DSIDLPHITPQFDGRSDYEAFILNGIPSGGLFTGAGIMSEENASRMGQAGVAYDANYH 404
DB 385 AAGRTSLPGPADGRSDYAVAFVLAGIPGLGSGFEAKTPQEAALMGTRGPRYDPCYH 444
OY 405 AAGDMTNLNEAFILNSKATAPAVATYA 433
DB 445 LTCDRTAQYSREAAATLHGQAFQTLTHYS 473
```

```

RESULT 3
OY 09Y646 PRELIMINARY: PRT: 472 AA.
AC 09Y646;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE AMINOPEPTIDASE.
GN Homo sapiens (Human).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RP TISSUE-LIVER.
RC TISSUE-LIVER.
RA LIO C.H., LIN B.Y., CHANG L.Y.;
RT "Cloning of the human aminopeptidase gene."
RL submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF107834; AAD43214.1; -.
KW Aminopeptidase.
SO SEQUENCE 472 AA; 51887 MM; A60A6716 CRC32;
```

Query Match 9.2%; Score 229.5; DB 4; Length 472;
Best Local Similarity 22.6%; Pred. No. 4.2e-08;
Matches 96; Conservative 83; Mismatches 175; Indels 71; Gaps 19;

```

OY 38 RVFGSKAHDDTVNLYELKTKGYDYDYVKOP-QVHLSNADOTLKVGDEIE-----AK 90
DB 77 RLSSKMLEKAIOIMYQLODQDEKYNLEPRVPHWREGESAVMLEPRHKIALILGLG 136
OY 91 TMTSPSVEVTAADVAVYVNNLGCSEADYPSDVEGKVALIKRGCEPFGDKS-----V 140
DB 137 SSGTPEPGITAEVLYVTSPELOR-RASEARKIIVYNNQ---PYINYSRVYRTGAV 192
OY 141 LAARAKAAASSTIVNNVNASMAGTLGAAG--SDKGPYSIVGISELEDGOKLIKLAAG-SV 197
DB 193 EAAVAGALASIRSVASPSISPTGIEQODGVKIPACTIVTEDAEMMSRMASHGKI 252
OY 198 SYDLMTWDSKOENRT-TYNVAQITGSD-PNNVYALGHTSVEAGPINDGSGSIISNLV 255
DB 253 VIQKMGATKTPDDSDSNTVAEITGSKYPRQVYLVSHLDSWVGQAMDDGGCAFISWE 312
OY 256 IAKLTQYSV--KNAVRFLEWTAEEFGLGS--NYVSHLNATELNKIRLYLNFMASPN 312
DB 313 ALSLIKDLGKPKRTLRVLTAEAGGVGAFQY-----QIAKYN-----SN 356
OY 313 YALNITDGDGSAFNOSGP--AGSNOIKLEFED---YYSDIDLPHITPQFDGRSDYEAFIL 367
```


RT "Cloning of the mouse aminopeptidase gene";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF07835; AAC3215.1; -
 KW Aminopeptidase.
 SQ SEQUENCE 470 AA; 51813 MW; 11BBD0F5 CRC32;

Query Match 8.6%; Score 215.5; DB 11; Length 470;
 Best Local Similarity 22.6%; Pred. No. 3.8e-07;
 Matches 96; Conservative 81; Mismatches 178; Indels 69; Gaps 18;

38 RVFGKAHDDVNTLYELKKTGYDYV-KQPVHMSNAQTLKVGDEIEAKTM---- 93
 75 RLSSKNEKAIQIMYQLQDDGLNVHLEQVRIPIHMERGESAVMLEPRHKKAILGLG 134
 93 --TSPSEVTVADVAVVKNLCCSEADYPSDVEKVALIKKECPFG-----DKSV 140
 135 SSIGTPGGITAEVLVVASFDELQR-RASEARGKIIVNQ--PYTGKTYQYRVQGVAV 190
 141 LAARAKAAASIVYNNVAGSMAGTIGAQ--SDKGYSAIVGISLEDGKLIKLAEGS-V 197
 191 EAAVGAASVLIQSVASISYSPHTGIQKIDVYKPIPTACTIVEDMMSSRMASRGKTI 250
 198 SYDLMDVSKOENRT-TYNNVAQTKG-DPNNVVALGHTDVEAGPGINDSGSIIIS-- 254
 251 VILHEMGKTYPTDTSFTVAETIGSMYPEEVVLVSGHLSMDVGGALDDGGAGFTSWE 310
 254 -LVAKALTOYSVKNVAFLEWTAEFGLLGSNTYVSHLNTLNKIRLYINFDIMASPN 312
 311 ALSLVKDL-GURPRRTLVLTAEQGGIGASQY-----ELHKANI-----SK 354
 313 YALAIYDGGSAFNOSGP--GSAOIEKLFEDYDIDLPHIPTQFDG--RSDYEAFILN 368
 355 YSL-VMEADSGTFLPTGIQFGSDKARAIKEVNNLQPLNVTAFVSGESTDINFTQA 413
 369 GIPSGGLFTGAIGMSEENASRMGQAGVAYDANYHAAGDMNTNLNHEAPLINSKATAFA 428
 414 GVPASLRDDL-----YKYEFFHSHSGDTMTVMDPRQ--MNVAAAVMA 454
 QY 429 VATY 432
 Db 455 VVAY 458

RESULT 7
 076552
 ID 076552 PRELIMINARY: PRT: 493 AA.
 AC 076552;
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
 DE AMINOPEPTIDASE ES-62 PRECURSOR.
 OS Acanthocheilichthys vitreus (Dipetalonema vitreus).
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Spirurida; Spirurida;
 OC Filarioidea; Onchocercidae; Acanthocheilichthys.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA HARNETT B., HOUSTON K.M., BRZESEKI H., APPEL H., ADAM R., TATE R.,
 RA GARATE T.;
 RT "Molecular cloning of, and assignment of aminopeptidase function to, a
 RT major filarial nematode secreted glycoprotein."
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF077194; AAC28365.1; -
 KW Signal; Aminopeptidase.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 493 AMINOPEPTIDASE ES-62.
 SQ SEQUENCE 493 AA; 54916 MW; B63093DF CRC32;

Query Match 8.3%; Score 207.5; DB 5; Length 493;
 Best Local Similarity 20.7%; Pred. No. 1.5e-06;
 Matches 107; Conservative 89; Mismatches 197; Indels 125; Gaps 22;

QY 5 VSPDEPEDI---QLBDELSQQLDEFAVYPERNRVSGKAH-----DDTVNYLY 53
 Db 28 VAPKNYIOETFGFEVALLIOTITGEEVGLAYOMLSLWDGFGHRMGSDLSKSTAFLE 87
 QY 54 ELKKTGYDYVYKQ--POVHMSNADQTLKVG--GDEIEAKTMYSPEVTVADVAVVK 108
 Db 88 ESLKNDNFQVHTEEVNPLPHWVRGNVDVEMIEPRNQLNVLAIGSEPSASATGETVIV 147
 QY 109 NLGCGSEADYPSDVEKVAL-----IKRGECPGDKSVLAARAKAAASIVYNN 155
 Db 148 DL---DVKRPDDVGRKIVTAQTFAGYPLTLKRYRSVKLFEOGALGVLYKSTISFSINS 204
 QY 156 V--AGSMAGTIGAQSDKPGYSIAVGISLEDGKLIKLAEGSVSYDLMDVSK--OEN 209
 Db 205 PHGTGAEINTTIRA-----CLITEEAMLERLYRSK-KIVYMDKSHYEE 252
 QY 210 RTYNNVAQTKGD-PNNVVALGHTDVEAGPGINDSGS--IISNLVAKALQ--- 263
 Db 253 INSNLIFEITGSRPEEVLLSAHVDSDVGGALDDGACAVMSALHSLKKLAERNP 312
 QY 263 -YSKKNVRFLEWTAEFGLLGS-NYVSHLNTLNKIRLYINFDIMASPNALAIYDG 320
 Db 313 KFRPRRTIRGIETSEQYGGAKHYITRKN-----DSPEKTFVSE 356
 QY 321 DGSAFNOSGPAQAQIEKLFEDYDIDLPHIPTQFDGSDY-----EAFILNGIPSG 374
 Db 357 DTGTFKSTN-----WLAHLSFSDSKSMRLKRETRLSNGIALGL 398
 QY 375 LFTGAGISMEENASRMG--GQAGVAYDAN-----YHAAGDMNTNLNHE--AFILN 421
 Db 399 INSNVQG---DVTEFAKDGIPSVNIPDKAVDYFFYFHHTGADYMTVLKDGDELYTTS 453
 QY 422 SKATFAVATYANDLSIP-----KRTTSSLHR 451
 Db 454 IFATLGHVIANMDWGSDFNQQLNSKOSTTEKSDRK 491

RESULT 8
 067732
 ID 067732 PRELIMINARY: PRT: 368 AA.
 AC 067732;
 DT 01-AUG-1998 (Tremblrel. 07, Created)
 DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
 DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
 DE HYPOTHEICAL 42.1 KD PROTEIN.
 GN AQ1891.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RX MEDLINE: 98196666.
 RA DECKERT G., WARREN P.V., GASTERLAND T., YOUNG W.G., LENOX A.L.,
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AJCAY M., HUBER R.,
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus."
 RL Nature 392:353-358(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RA DECKERT G., WARREN P.V., GASTERLAND T., YOUNG W.G., LENOX A.L.,
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AJCAY M., HUBER R.,
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF000762; AAC07705.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 368 AA; 42078 MW; C6BFC83 CRC32;

Query Match 7.6%; Score 189; DB 2; Length 368;
 Best Local Similarity 23.4%; Pred. No. 1.7e-05;
 Matches 79; Conservative 66; Mismatches 147; Indels 46; Gaps 13;

```

QY 36 RNRVGGKADDTVNTLYEELKTKGYDYVYKOP-QYHLMNSNADQTLKVGDEIEIAKTMTY 94
DB 19 KNRISGTKNKKARRKFKNNYLRKRG-FKYEDFSTYKTPVSARIKCEGEFPAPLLIG 77
QY 95 SSVSEVTAADVAVVNNKLGSEADYPSDVEGKVALIKRGCEPFGDKSVLAAKAKAAIYVNN 154
DB 78 SLMGSTSGEVVILKNNL-----EKKDLKGVVALPVGGRDSEKAFLEDKRAAGMT- 132
QY 155 NVAGSMAGTLGAAGSDKGYSAIVG-----ISLEDGOKLIKLAEGSVSDI,AVDSKO-EN 209
DB 132 -----LELDVPSGSGIDGVKFFAVNTREIYKRIEKGKVVLD:RREKKEING 179
QY 210 RTTYNVVAOTKRGDNNVVALGHTDSVAGPGINDDGGIISNLYIAKALQY-SVKNA 268
DB 180 KNLV--VERGRG---PVLLVAHYDTKPPYGAIDNGLSVALLLVMSREL,AGFOEIPPR 233
QY 269 VFFLEFWTAEEFGILGNSYVSHLNATELNKIRLYNFDMIASPNVALMIDYDGGSAFNOS 328
DB 234 IILITDCEDELLEGASHALH-----SKNIFYVIMLDSIGKNRAVLVEDA-----S 281
QY 339 GPAGSAQIEKLFEDYD-SIDLPPIPTQFGSDYEAF 365
DB 282 GENGRLINEKLFKHLMDKMDIPVEESK-TGMSDHVPE 318

RESULT 9
P70627 PRELIMINARY; PRT; 752 AA.
AC P70627;
DT 01-FEB-1997 (TREMblrel. 02, Created)
DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE NAAG-PEPTIDASE (GLUTAMATE CARBOXYPEPTIDASE II) (EC 3.4.17.21).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAQUE DAWLEY; TISSUE-BRAIN;
RX MEDLINE; 98041505.
RA BZDEGA T., TURTI T., WROBLEWSKA B., SHE D., CHUNG H.S., KIM H.,
NEALE J.H.;
RT "Molecular cloning of a peptidase against N-acetylaspartylglutamate
from a rat hippocampal cDNA library.";
RL J. Neurochem. 69:2270-2277(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE; 98169524.
RA LUTHI-CARRIER R., BERGER U.V., BARCZAK A.K., ENNA M., COYLE J.T.;
RT "Isolation and expression of a rat brain cDNA encoding glutamate
carboxypeptidase II.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:3215-3220(1998).
RN [3]
RP SEQUENCE OF 284-752 FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE; 96149377.
RA "Prostate-specific membrane antigen is a hydrolyase with substrate and
RT pharmacologic characteristics of a neuropeptidase.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:749-753(1996).
RN [4]
RP SEQUENCE OF 284-752 FROM N.A.
RC TISSUE-BRAIN;
RA LUTHI-CARRIER R., BERGER U.V., BARCZAK A.K., ENNA M., COYLE J.T.;
RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; U75973; AAC53423.1;
DR EMBL; AF040256; AAC40067.1;
DR EMBL; AF039707; AAB96759.1;
KW Carboxypeptidase; Hydrolyase.
SQ SEQUENCE 752 AA; 84539 MW; 9C4B29D8 CRC32;

```

```

Query Match 6.6%; Score 166; DB 11; Length 752;
Best Local Similarity 20.4%; Pred. No. 0.002;
Matches 97; Conservative 61; Mismatches 151; Indels 166; Gaps 22;

QY 18 DLLESGQLEDAVAVYNNRNVFGKADDTVNTLYE-----ELKTKGYDYK 66
DB 108 DVEISDY-DVLLSTP-----NKTHPNYISTINEGNEFTKSLAEISPEIGENT-- 157
QY 67 QPQVHLMNSNADQTLKVGDEIEIAKTMYSPEVETADVAVYK-----NLGC 112
DB 157 -----SDVPPYSAFSPGTPGDLVYVYARTDEDFKLEBYVKINC 198
QY 113 SEADYPSDVEGKVALIKRGCEPFGDKSVLAAKAKAAIYVNNVA----- 158
DB 199 S-----GKIVIARYGQYRGNKVNQAQAGKGIILYSDPADYFVPGKSPDGN 249
QY 158 -----GSMAGTLGAAG-----SDKGYSAIVG-----ISLEDGOKLIK 190
DB 250 LPGGVQGRNVNTLNGAGDPLTPGPANETAYRHEFTETAVGLPSIPVHIGYDQAKLIE 309
QY 191 LAEGSVSYDLVWD-----SKOE-----NRTT--YNNVAOTKRG-D 223
DB 310 HNGGSAAPDSSMKGLKPYNVPGFAGNFSKQVKLHHSYNTVRYNVIIGTLKAVE 369
QY 224 PNNVVALGHTDSVAGPGIN-DDSGGIISNLYIAKALQ--YVKNAVRLEFWTAEEF 279
DB 370 PDYVILGGHRDAWVFG-GIDPQGAAYVHEVTFGTLLKKKMPRRITLESADAEEF 428
QY 280 GLGGSNYVSHLNATELNKIRLYNFDMIASPNVAL-----MIDYDGGSAFNOS 329
DB 429 GILGSTEWAEEHSLQGERGVAYINADSSIEGNTLYLRDCTPLMSIYNN-----LTREL 483
QY 330 PAGAQAIEKLFEDYDS-----IDLPHIPTQFGSDYEAFILN-GIPSG 373
DB 484 PPDGDFEG--KSLYDSWKESPSSTFIQMPRI-SKLSGNDFEVFFORLGIASG 535

RESULT 10
P70216 PRELIMINARY; PRT; 433 AA.
AC P70216;
DT 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
DT 01-AUG-1998 (TREMblrel. 07, Last annotation update)
DE HEKATOPOIETIC LINEAGE SWITCH 2.
GN HLS2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA WILLIAMS J.H., CHAN C.-Y., KLINKEN S.P.;
RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF009513; AAC17945.1;
SQ SEQUENCE 433 AA; 47858 MW; 655ACEF9 CRC32;

Query Match 6.6%; Score 164.5; DB 11; Length 433;
Best Local Similarity 21.3%; Pred. No. 0.0011;
Matches 86; Conservative 76; Mismatches 163; Indels 79; Gaps 18;

QY 38 RVFGKADDTVNTLYEELKTKGYDYV-KOPQVHLMNSNADQTLKVGDEIEIAKTM-- 93
DB 75 RLISGKNKEKAQIOMYONLODGLENVHL,EOYRIPHWERGESAVMLPRIHKMAILGIG 134
QY 93 --TSPSEVTAADVAVVNNKLGSEADYPSDVEGKVALIKRGCEPFG-----DKSV 140
DB 135 SSTGTGPGITAEVIVVASFDELQR-RASEARGKTIIVNQ---PTTGTEKYVQYVQCAV 190
QY 141 IAAKAKAAIYVNNVAGSMAGTLGAAG--SDKGYSAIVGISLEDGOKLIKLAEGSV-V 197
DB 191 EAAKGAVAASLIQSAFSPISPHTGIOYQDGVKPIPTACTIVTADAEWMSMAARGNKI 250

```

Oy	198	SVDLVADSKOEKRR-TYVNVVQOTGG--DPNNVVALGHTDSEVAGPGINDGSGISN--	254
Dd	251	VHLEMGKATTPDDISFTVAETIGSMPEVY-----GQALDDGGAFISWE	298
Oy	254	-LVIAKALTOYSVKNVAVRFLFWIAEEFGLSGSNYYVSHLNTLNLKIRLYLNFDMIASPN	312
Dd	299	ALSLVKDL-GLRPKRTLRLVMTFAEEGGICASQY-----ELHKANI-----SK	342
Oy	313	YALMTYDDGSAFNOSGP--AGSAQIEKLFEDYDSDLDPIRPIQFGG--RSDEAFILN	368
Dd	343	YSL-VMEADSGTFPLPTGQFTGDSKARAIKMEVNNLLQPLVTVYFVNGEGTDINFIQA	401
Oy	369	GIPSGLFTGAEGIMSEENASRMGQAGVAVDANYHAAGDMNTN	412
Dd	402	GVPGASLDDL-----YKIFFHHSHGDMTIS	428
RESULT	11		
P91406			
AC	P91406	PRELIMINARY;	PRT; 751 AA.
DT	01-MAY-1997 (TrEMBLrel. 03, Created)		
DT	01-MAY-1997 (TrEMBLrel. 03, Last sequence update)		
DT	01-NOV-1998 (TrEMBLrel. 08, last annotation update)		
DE	SIMILAR TO HUMAN PROSTRATE-SPECIFIC MEMBRANE ANTIGEN.		
GN	R57.1.		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;		
OC	Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.		
RN	[1]		
RN	SEQUENCE FROM N.A.		
RC	STRAIN-BRISTOL N2;		
RX	MEDLINE; 94150718.		
RA	WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,		
RA	BROFIELD J., BUTTON J., CONNELL M., COSEY T., COOPER J., COULSON A.,		
RA	CRAWFORD M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,		
RA	GADNER A., GREEN P., HAWKINS T., HILLIER L., JIR M., JOHNSTON L.,		
RA	JONES M., KERSHAW J., KIRSTEN J., LAISER N., LARRELLE P.,		
RA	LIGHTNING J., LOYD C., MEMURRAY A., MORTIMORE B., O'CALLAGHAN M.,		
RA	PARSONS J., PERCY C., RIFEEN L., ROOPER A., SAUNDERS D., SHOWNKEEN R.		
RA	SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,		
RA	THIERY-MEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,		
RA	WATSON A., WEINSTOCK L., WILKINSON-SPRAT J., WORLDMAN P.;		
RT	"2.2 Mb of contiguous nucleotide sequence from chromosome III of C.		
RT	elegans."		
RL	Nature 368:32-38(1994).		
RN	[2]		
RN	SEQUENCE FROM N.A.		
RC	STRAIN-BRISTOL N2;		
RA	FAVELLO T., RIFKIN L., CHIARELLI B.;		
RL	Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RN	SEQUENCE FROM N.A.		
RC	STRAIN-BRISTOL N2;		
RA	WATERSTON R.;		
RL	Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-BRISTOL N2;		
RA	WATERSTON R.;		
RL	Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.		
QC	EMBL; U88179; AAB52660.1; --		
QC	SEQUENCE 751 AA; 83683 MW; 088B6F82 CRC32;		

QY 78 QTL---KVGDEIEEK-----TMTSPVEVTAADAVYKNGCCSA 115
Db 115 SVITTSAGKEVEFKSKGVSPVLIIPDQSKYAGHOMLAYAGNGSADAVYIT-NHGTAN- 173
QY 116 DYPS-----DVEGRVALIKRGECEFGDKSVLAKAKAAASIVYNNV----- 157
Db 173 DEKNLKLGVDPDIKGTALIMRYGHRGRDKIHKAQOAGAILFSDTQVADQGVSENV 232
QY 157 -----AGSMAGTLGAQSD-----KG-----PISAIYGISL 182
Db 233 YPKIIMPNNEGVOQRSLMHGDGDLSPYPPSKKELEKGTIEAKEDGVLPSIPVLPVS 292
QY 183 EDGOKLIKLAEGSVSYDLM-----VDSKQENRTTYNVA 217
Db 293 TTGYITLILKLSRRPASP-WQGFVVGGNLTIKAGPFGVNGEKISINVHSLPRKRIIRNVIG 351
QY 218 QTKGG-DEPNNVYALGGHTDYSVAPGINDGSGIISNLIAKAL-----TOYSYKNAV 270
Db 352 YLRGEEDPSYIMLNHFDAWYGSIDPNSGTAVALAE-VARAAQGTINETSMPKPARTIV 409
QY 271 FLEFWAEEFGLGSGNYVYSHLNATELNKRIYLNFDMLASPNYALMTYGDGSAENQSCP 330
Db 410 FNMAAEERGLGSTEFVEEFYNILOKRAVYIINDCI-----OGNISLHYDTVP 459
QY 331 -----AGSAQIE-----KLFEDYDSIDLPHIPTQDGRSDYA 364
Db 460 ILEHVAIENSKOVENPSKRSRGRKGLTYDTMKKFFPKKAGVPKIRV---GGGSDHAP 516
QY 365 FILNGIPSGGLETGAEGI-MSEENASRWGGQAGVAYDANYHAAGDNMTNLNHEAPLIN-- 422
Db 517 F-LN-----FAGVPVINFTEFKNYTTW--DYPLHYHTXETPEFSIHLLDIDNLVSKA 566
QY 422 -SKATAFAVAITANDLSSIPKRNITSSILHRRARTKRPPEKRAPKTHAHVSGGCHSQVE 480
Db 567 IGVYAEALAKTFADV--ILPMNT--HFASVMLKTY---LPQKTTISGINVSRSDPE 618

RESULT 12
057813 PRELIMINARY; PRT; 565 AA.
057813:
ID 057813
AC 057813
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)
DE 565AA LONG HYPOTHEITICAL PROTEIN.
GN PH0048.
OC Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
[1]
RC SEQUENCE FROM N.A.
RC STRAIN-073;
RX MEDLINE; 98344137.
RX KAWARABAYASI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,
RYAAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOYAMA A., NAGAI Y.,
RA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMITA M., OHFUKU Y.,
RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAOKI J., KUSHIDA N., OGUCHI A.,
RA AOKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,
RA KIRUCHI H.;
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.",
RL DNA Res. 5:55-76(1998).
QO EMBL, AF000001; BAA29116.1;-
QO SEQUENCE 565 AA; 63745 MW; 4CB63039 CRC32;

Query Match 6.5% Score 163 DB 5 Length 751;
 Best Local Similarity 19.0% Pred. No. 0.0032;
 Matches 114, Conservative 91, Mismatches 21, Indels 178, Gaps 26.

23 SQQLEDFYATPENNRRVGGKAHDDTVNLYEELKKTGYDVKQPPQVHLMS----NAD 77
 :::::|||||
 55 TENIKRYRITFKDPAVAGTEANKRVAYEIAANWSEAGDEADVETLPYEYELLVDPDEFEN 114
 :::::|||||

```

Query Match      5.3%; Score 158; DB 1; Length 565;
Best Local Similarity 21.8%; Pred. No 0.0046;
Matches 104; Conservative 78; Mismatches 173; Indels 122; Gaps 28;

QY 4 LVSPDEF-PEDIQLEDLIEGSOQLDEFAYVPERNRVFGKRAHDTVNYLYEELKKTG-- 61
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
12 LKADKFNPEV--FNDVVEIS-----KFRIGSRDVRVATEVYVSRLELGD 59

```

```

QY 61 ---YDYKQPOVHL-----WGNADOTLKVGDEIEAK-----TMTYPSVEVTADVA 105
D 60 YKFLNDIYDGKSYHLTLPSPIGWEIVGSLKFKKEKELTNSPLLVAHNSPGSAKGEVI 119
QY 106 VKNLGGSEADYPSDVEGKVALIKRGCPFGDKSVLAIAKAK---AAASIYNNVAGSMAG 162
D 120 PI---FREEDW-EKADGKILV--GE---DMREAYRKANEGHAKAFIAYRKGTGDAFP 168
QY 163 TLGAAGSDKG-PYSAIVGISLED--GOKLIKLAAGSVYDLWVDSKOENRTYNNVAQT 219
D 169 YIGFLFKKOLEMAKIALTYPEVANLIEKSKKGVEVSVKTEIKDEIPL-VA 227
QY 220 KGGPNNVVALGHTDVEAGPGINDSGGILSLVIAKALTQYVKNVAFWLWAEF 279
D 228 KIGPPPL-L-PSAH-L-CHPKRGANDNAGSGAMLIELARILKNEGRIQPAFL-WIDEPH 283
QY 280 GLGASNYVSHLNTLTKIRLYLNFPMYIAS---PNYALMI----- 318
D 284 ---GTQAFIPRYN---LNEITANINLDVYGSGEDRSNSTIMLVRTPLUSLEVPGLLEF 337
QY 318 ---YDGDGSAFNOSGAPGSAQIEKLFEDYDIDLPHPITQDGRSDYPAFILNGIPSGG 374
D 338 LKEVNAEGKSPSGS-PLPSVK-----PKFY-----PYEKGSDHDIENFYSIP--- 379
QY 375 LFTGAEIGMEENASRNGGAGVAYDANYHAAGDMNTLNHEALINSKATAFAVAT 431
D 379 ---GVM---PITWP-----DRYHTSADTPEKLSLRLSLIGRA---VYAT 445

RESULT 13
077564 PRELIMINARY; PRT; 751 AA.
ID 077564;
AC 077564;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE POLY(POLY-GAMMA-GLUTAMATE CARBOXYPEPTIDASE.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=JEJUNAL MUCOSA;
RX MEDLINE; 98352082.
RA HALSTED C.H., LING E., LUTHI-CARTER R., VILLANUEVA J.A., GARDNER J.M.,
RA COYLE J.T.;
RT "Poly(poly-gamma-glutamate carboxypeptidase from pig jejunum.
RT Molecular characterization and relation to glutamate carboxypeptidase
RT II."
RL J. Biol. Chem. 273:20417-20424(1998).
DR EMBL; AF050502; AAC39269.1; -.
KW Carboxypeptidase.
SQ SEQUENCE 751 AA; 84523 MW; 0F6E680A CRC32;

Query Match 6.3%; Score 158; DB 6; Length 751;
Best Local Similarity 19.7%; Pred. No. 0.0071;
Matches 98; Conservative 72; Mismatches 188; Indels 140; Gaps 20;
QY 3 ALVSPDEFPEIDQEDLEGSOQLEDFAYAVPERNRVFGKADHDTVNYLYEELKKTG-- 61
D 50 ANISPOHNNKAFIDEL--KAENIKTLFLYNTRIPHLAGTQNLAKQIOSQKKEGLD 107
QY 61 ---YDYVKKQPOVHLMSNADQT---LKVGE-----EIEA 89
D 108 SVELAHADV-----LLSYNPKTRPNYISIIDGNEIFNTSLFEPPEPEYENVSVP 160
QY 90 KTMYYSSVEVTADVAVY---KNLGCSEADYPSDVEGKVALIKRGCPFGDKSVLAIAK 144
D 161 PFSAFSGMEGDLVYVNVARTEDFKLERDKMINKSGKILLIARYGIFNGKNVKNQOL 220
QY 145 AKAAASIVYNNVA-----GSMAGTLGAAGSDKGPYSA----- 177

```

```

D 221 AGAKGILISDPADYFAPVQSYPDGWNLDGGVQKQNNILNAGAGDPLTPGIPANEYAT 280
QY 177 ---IVG-----ISLEDGOKLIKLAAGSVSYDLN----- 203
D 281 RLQIAEAVGLPRIPVHIGYSDAQKLEKMGSGAPPPDSKSGSLHVPYNGPFTGNFST 340
QY 203 ---VDSKOENRTYNNVAQTKG--DPNNVVALGHTDVEAGPGIN--DGGSGIISNL 234
D 341 QKVAMHSHSNKVKRYINVTGLGAVEPDRYVILGGHRDSWFG--GIDPGSAAVYHEI 399
QY 255 V--IAKALLO-YVKNVRFLEWTAEFEGLGSNYYVSHLNTLTKIRLYLNFMDIASP 311
D 400 VRSFKLKKGWRPRRTVLPSWDAEYGLFGSTEMEENSRILQEKVAYINNDSIEG 439
QY 312 NYAL-----MIDGDSAFNOSGAPGSAQIEKLFEDYDIDLPH---IP--TQF 356
D 460 NYLRFVDCFTPLMSLYVN---LTKELQSPDEGEKSLFESWMEKSPSPFGSLPRISKL 516
QY 357 DGRSDYAFILN-GIPSG 373
D 517 GSGNDFEVFFORLGIASG 534

RESULT 14
082996 PRELIMINARY; PRT; 393 AA.
ID 082996;
AC 082996;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 12, Last annotation update)
DE AMINOPEPTIDASE.
OS Aeromonas caviae.
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonas group;
OC Aeromonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-T-64.
RA HAYASHI K., IZAWA N., NIRASAWA S.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-T-64.
RA IZAWA N., HAYASHI K.;
RT Cloning and nucleotide sequencing of the aminopeptidase gene from
RT Aeromonas caviae T-64."
RL J. Ferment. Bioeng. 82:544-548(1996).
DR HSSP; Q01693; 11GB.
KW Aminopeptidase.
SQ SEQUENCE 393 AA; 42228 MW; C6CB958C CRC32;

Query Match 6.3%; Score 157.5; DB 2; Length 393;
Best Local Similarity 25.2%; Pred. No. 0.0028;
Matches 78; Conservative 37; Mismatches 113; Indels 81; Gaps 17;
QY 158 GSMAGTLGAAGSDKGP-YSAIVGISLED--GOKLIKLAAGSVYDL-WVDSKOENRTT 212
D 123 GNIYGRISQLASMRNRYTTTGTGVSADWVAGQ-----WQSLSATLPLWASVSRVHSG 175
QY 213 Y---NYVAQTKGQD-PNNVVALGHTDVSVA-----GPGINDSGGILSLVIAKAL 260
D 176 YPOOSVYLITLKGRKYPDEVVVLGGHLDSTAGSAPNSRTIAPGADDDASGIATLEVLRY 235
QY 261 TOY--SVKNVRFLEWTAEFEGLGSNYYVSHLNTLTKIRLYLNFMDIASPNVLMY 318
D 236 AEOGRPERLQFTIGVAAEVLGRSKDITRYKAAN-TVYLAALQIDM---TNY----- 287
QY 319 DGDGSAFNOSGAPGSAQIEKLFEDYDYS-----ID--LPHIPTQFD---GRSDYE 363
D 287 ---GSAADIYFMTDYTPQGLTGVLQALDLYLPQIRYGYDSCGCGSDHA 334

```


Thu May 11 10:03:21 2000

us-09-080-127-2_copy_16_496.rspt

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2000, 21:15:26 ; Search time 984.82 seconds
(without alignments)
-1425.375 Million cell updates/sec

Title: us-09-080-127-1_COPY_46_1488
Perfect score: 1443
Sequence: 1 GGAGGGCGCTGTTCGCC.....GGCATTCACAGTCAGAGCA 1443

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 segs, -486395729 residues

Total number of hits satisfying chosen parameters: 176538

Minimum DB seq length: 0
Maximum DB seq length: 100000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ov:*
4: gb_ov:*
5: gb_ov:*
6: gb_ph:*
7: gb_ph:*
8: gb_ph:*
9: gb_ph:*
10: gb_ph:*
11: gb_ph:*
12: gb_ph:*
13: gb_ph:*
14: gb_ph:*
15: gb_ph:*
16: gb_ph:*
17: gb_ph:*
18: gb_ph:*
19: gb_ph:*
20: gb_ph:*
21: gb_ph:*
22: gb_ph:*
23: gb_ph:*
24: gb_ph:*
25: gb_ph:*
26: gb_ph:*
27: gb_ph:*
28: gb_ph:*
29: gb_ph:*
30: gb_ph:*
31: gb_ph:*
32: gb_ph:*
33: gb_ph:*
34: gb_ph:*
35: gb_ph:*
36: gb_ph:*
37: gb_ph:*
38: gb_ph:*
39: gb_ph:*
40: gb_ph:*
41: gb_ph:*
42: gb_ph:*
43: gb_ph:*
44: gb_ph:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	178.2	12.3	35420	1	MTG22G10
2	163.8	11.4	2272	7	YSCAM
3	163.8	11.4	2532	7	SCYBR286W
4	163.8	11.4	32421	7	SCDPB3
5	136.4	9.5	9120	1	SAPURCLUS
6	59.6	4.1	3480	1	SLAMC
7	57.2	4.0	42210	1	SCIC2
8	52.4	3.6	16911	1	SCE7
9	49.2	3.4	749	16	NPOEPT2
10	49.2	3.4	3932	7	ATAPG
11	49.2	3.4	35824	1	SCHE9
12	48.2	3.3	1207	8	AF130975
13	47.8	3.3	36368	1	SC985
14	47.6	3.3	92710	50	AC022472
15	47.6	3.3	131993	16	OPU75930
16	47.4	3.3	30001	5	AR031772
17	47.4	3.3	30001	5	I32939
18	47.2	3.3	64957	1	AB032367
19	46.8	3.2	64957	1	AB032367
20	46.8	3.2	97015	1	BSEGR
21	46.8	3.2	212150	1	BSUB0020
22	46.4	3.2	7836	2	AF047717
23	46.2	3.2	833	5	A37833
24	46.2	3.2	3573	1	SPU21215
25	46.2	3.2	5392	5	A37831
26	46.2	3.2	36850	1	MTG2369
27	45.6	3.2	2426	44	AC020448
28	45.6	3.2	47852	1	MTV023
29	45.6	3.1	42210	1	SCIC2
30	45.4	3.1	42210	1	SCIC2
31	45.2	3.1	740	7	S47405
32	45.2	3.1	152237	8	AP001073
33	45.2	3.1	38400	1	SC4H2
34	44.8	3.1	8018	2	AF164960
35	44.8	3.1	37200	1	SCE63
36	44.6	3.1	4761	1	SC0244019
37	44.6	3.1	32059	44	AC019906
38	44.6	3.1	36917	1	SC4G6
39	44.6	3.1	41816	2	SC7A8
40	44.6	3.1	45396	1	SCH35
41	44.6	3.1	68476	54	AC009200
42	44.4	3.1	11831	11	HS989H11
43	44.4	3.1	4462	2	AF126720
44	44.4	3.1	40476	2	SC8F4
45	44.2	3.1	3088	1	SSU08602

ALIGNMENTS

```

RESULT 1
MTCY22G10 35420 bp DNA BCT 17-JUN-1998
LOCUS Mycobacterium tuberculosis H37Rv complete genome; segment 21/162.
DEFINITION 284724 AL223456
ACCESSION 284724.1 GI:3261708
VERSION
KEYWORDS
SOURCE
ORGANISM Mycobacterium tuberculosis.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacteriales; Mycobacteriaceae;
Mycobacterium.
1 (bases 1 to 35420)
REFERENCE
AUTHORS Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C.,
Harris, D., Gordon, S.V., Eigmeier, K., Gas, S., Barry III, C.E.,
Tekala, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T.,
Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S.,
Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J.,
Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A.,
Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, S.,
Squires, S., Squares, R., Sultson, J.E., Taylor, K., Whitehead, S. and
Barrall, B.G.
JOURNAL Deciphering the biology of Mycobacterium tuberculosis from the
MEDLINE complete genome sequence
REMARK Nature 393 (6685), 537-544 (1998)
TITLE Erratum: [[published erratum appears in Nature 1998 Nov
12,396(6707):190]]
2 (bases 1 to 35420)
Parkhill, J.
REFERENCE
AUTHORS Direct Submission
TITLE Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium
JOURNAL tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
Trust Genome Campus, Hinxton, Cambridgeshire CB10 1SA Unité de Génétique
Moléculaire Bactérienne, Institut Pasteur, 28 rue du Docteur Roux,
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
On Jun 27, 1998 this sequence version replaced gi:1817670.
Notes:
Details of M. tuberculosis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/M.tuberculosis/) CDS have
been renumbered from the original cosmid submissions but the old
gene designations are in brackets after the new gene numbers.
Gene prediction was based on a Hidden Markov Model of TB genes
implemented in TParse (Krogh) supplemented with visual inspection
of positional base preference in codons, especially where there is
an increase in the observed/expected third position G + C.
CAUTION: In some cases we may not have predicted the correct
initiation codon. Where possible we choose an initiation codon
(ATG, GTG, or TTG) which is preceded by an upstream ribosome
binding site sequence (optimally 5-13bp before the most upstream
codon). If this cannot be identified we choose the most upstream
initiation codon.
FEATURES
source
location/Qualifiers
1. 35420
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
<1. 34666
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
/clone="Y22G10"
61..1071
/gene="Rv0407"
61..1071
/gene="Rv0407"
CDS
/gene="Rv0407"
/note="Rv0407, (MTCY22G10.03), len: 336, unknown coenzyme
F420-dependent enzyme, but similar to AEF01022-18
Archaeoglobus fulgidus SP1029071
N5,N10-METHYLENETETRAHYDROMETHANOPTER (348 aa); FASTA
scores: opt: 430 z-score: 514.7 E(): 3.8e-21, 28.6%

```

```

identity in 350 aa overlap; also similarity to hypothetical
37.3 kd protein MTCY02B10.24 (340 aa, 28.9% identity in
228 aa overlap) and to MTC15_6 (36.7% identity in 330
aa overlap)
/codon_start=1
/transl_table=1
/product="hypothetical protein Rv0407"
/protein_id="CAB06577.1"
/db_xref="GI:1817673"
/db_xref="SPRMBL:P66253"
/translation="MAELKLGKASAEQAFRELVELVAVAEAHGMDATVSDHFQW
RHOGGHAFPSMWTAVGERNRRLIGTSVLTPTFRNPATVIAQAFMGLPVNRYF
LGVETGELALNIEATGEGAMPEKEREKRLAEISYGLRMQLSGRVPEDGYLTKGA
SIYVPGDGPVYIAAGSPAYAKYAGAGAGDEICTSGKEELTEKEMLPVBECAAA
DESYDGIDKMEIKRISYDPEELNMTREAPRSTAEOKHSIDDEIEKADALP
IEQAKRRIYASDDEAVEKGVQYVTGMLNHLVHARGHDDRELPEFQSDIAPRLR
LG"
1064..3136
/gene="pta"
1064..3136
/gene="pta"
/note="Rv0408, (MTCY22G10.04), len: 690, pta, similar to
eg PTA_ECOLI P39184, phosphate acetyltransferase (713 aa),
fasta scores, opt: 1303, E(): 0, (38.0% identity in 718 aa
overlap)"
/codon_start=1
/transl_table=1
/product="pta"
/protein_id="CAB06578.1"
/db_xref="GI:1817674"
/db_xref="SPRMBL:P66254"
/translation="MADSAIYLAAPESQGTSTIALGLRLRLTAMAKYVERPITR
LSARDYLIELLAHTSAGLYRECVGTQOOLADDDAIAEIVDSYHMADECDYV
VYVSDYTDVTSPTRELNVNGRIANVGLAPVLTARADRPDOVASVEYVCLIEDVQ
RAHRAAYANRCELSATPVATDARTRTPSYVPEPLLSAPVLAELTAQVAGVYS
GDVALRREVGVAAGMTADHVERLTDGAVTPEDRSDDVVAVASAAGSPFSL
CSYALNGFOLHPAIAALVSGRLRLVATVATAGTYDTASAASARGLVATGDSQKID
TALIMDRHVDAQLALQLTIPPTVTPOMFTYRLQOARSDIMRFLVEGDDRLI
KSAGRLORGIIVDTIIGDEAKVRLRAEIGVDVAGATVEPCASELHDOYALD
RKAGTVEHAREITMDATFEGTMLVNHCHADGVSGAATHTATVPALEIKTVG
ISTYSIFIMKLPRIYATGDCATIPPTVBOQADLITCARTRAAQGTIPRANLST
STGSPGAGADVTKRAATELVRAEPOLPYEGPIQYDAAEPSYAAKLDSDPAVTA
TVLFPDLNNGNNTYKAVORSGAIAIGVQLGRKPNVLSRGALVDVIVNVAITA
IOAGQVHE"
3129..4286
/gene="ackA"
3129..4286
/gene="ackA"
/note="Rv0409, (MTCY22G10.05), len: 385, ackA, acetate
kinase, similar to eg ACKA_BACSU P37877 acetate kinase
(ec 2.7.2.1) (395 aa), fasta scores, opt: 974, E(): 0,
(43.5% identity in 393 aa overlap), contains P801075
Acetate and butyrate kinases family signature 1, P800758
ArgE / dape / Act1 / Cpg2 / yscs family signature 1"
/codon_start=1
/transl_table=1
/product="ackA"
/protein_id="CAB06579.1"
/db_xref="GI:1817675"
/db_xref="SPRMBL:P66255"
/translation="MSSTVAVINSGLSLFOVPEVAGNSRAGIVERIGNSPVA
DHAQLHRAFKMLAEDGIDLCGLVAGRVVGGTGPPTLDDTVIGKEELSA
LAPLHPNPAVIGIKVARRLANVAVPTAFPHDPPAAATYAIADRDVAHRIIR
XGFGTSHOVYSEERAAPLGRPLDGLNOIYLHNGSASAIAGRREVSMTGPTPE
GLVNGTRSGDLDPGVISYLMRTAMGVEDIESMHNHSGMLGAGEDFRRLRYVET
GDRSAQLAYEVFHLRLKRYLGATYAVAGHDVVSFTAGIENDAVRDADLAGQGLG
IALDQDRNLGPGHGARRISSDPSIAVLVPTVEELAIADCLVILGGRRA"
3141..3176
/gene="ackA"
/note="P801075 Acetate and butyrate kinases family
signature 1"
4026..4055
/gene="ackA"

```

misc-feature

```

/Note="PS00758 Arge / dape / ACY1 / CP62 / yscs family
signature 1"
gene complement(4340..6592)
/ gene="pkng"
/ complement(4340..6592)
/ note="Rv0410c, (MTCY22G10.06c), len: 750 aa, protein
kinase similar eg to PKNA_MYCLE P54743 protein kinase
pkna, mycobacterium le (253 aa), fasta scores, opt: 291,
E(): 3.1e-10, (33.2% identity in 203 aa overlap); contains
PS00108 Serine/Threonine protein kinases active-site
signature"
/ codon_start=1
/ transl_table=11
/ product="pkng"
/ protein_id="CAB06580.1"
/ db_xref="GI:1817676"
/ db_xref="SPTREMBL:P96256"
/ translation="MAKASETERSGPGTQPADQATATATVRLSTQAVRPPDGED
NRPPLTGPDPEDQDMATTSRYRPPRYRLGGGLVETPRAPDIPFALMTNPVYES
KRFQWNGRPGVSDSETKASGEGWCPGSPISFQPLNPGDIVAQVEKCIAGS
GLGWITLADNRNNGRPVYKGLVHSGDAEQAMAAERQFLAEVHPSPVQVFNFE
HIDRHGDPVGYIYMEYVGOSLRSKSGOKLPEAALAYLEILIPALSHSISGLVND
LRPENIMLFEOLKIDLGAVSRINSFGYLGPGQAPVIRGPTVADITVGRNT
LAATIDLPTRNGRYVDGLPEDDPVLTQDYSGRILRAIDPDPRQRTTAERMSOL
TGLREVVAADITVPRPGSTSPRSSTFGVDLVAHNDVYIDGQVHAKEKLTANETV
TALSVPVDPDVAASVLDATVLSQVQTLDSLRARHGLDLDGVDSESEVPLME
VRALDLGVAKATRKLDLDERVGRWRLVWRAVAELTGTGDIHETFEVDLTFE
PGEIARLALAAELAGNTEDEHEFYQVSTWSTGDIYSAFGLARARSAGDVGAIVR
TIDVEVPTSRHETARTSATVLSGSRSEVTEQIRDAARVEALPPEPVQIDR
ALVIGALMDIMKONKASTNHILGFPFTHGLRIGVASLSRLARVAPLQRHRTLVDM
ANKVRSPTSF"
misc_feature complement(5741..5779)
/ gene="pkng"
/ note="PS00108 Serine/Threonine protein kinases
active-site signature"
/ complement(6592..7578)
/ gene="glnH"
/ complement(6592..7578)
gene / gene="glnH"
/ note="Rv0411c, (MTCY22G10.07c), len: 328 aa, glnH,
similar to eg GLNH_BACST P27676 glutamine-binding protein
precursor(262 aa), fasta scores, opt: 493, E(): 7.5e-22,
(37.8% identity in 193 aa overlap), contains PS00013
Prokaryotic membrane lipoprotein lipid attachment site,
PS01039 Bacterial extracellular solute-binding proteins,
family 3 signature"
/ codon_start=1
/ transl_table=11
/ product="glnH"
/ protein_id="CAB06581.1"
/ db_xref="GI:1817677"
/ db_xref="SPTREMBL:P96257"
/ translation="MTRRALIARAAPLALAVLAVASGHSSETLVEATPLPPT
PVGMEIMPPOPPLPPDSSODDPTASLRPAKRAADAADNIRARGILVIGLIDGS
NLSFSDPTGTGTVGVDIAGVARDICGVPRVYHRTLSAERTALOKSDVDIVY
KIMSTICERKLVNFTVYLDANQRIARPDSFTTVSDISGRKVCVAAKGTISLRNIR

```

```

Query Match 12.3%; Score 178.2; DB 1; Length 35420;
Best Local Similarity 56.2%; Pred. No. 1.6e-24;
Matches 386; Conservative 0; Mismatches 283; Indels 18; Gaps 2;

QY 640 AACGTTTCGCCGACGAGGAGGCGGCGGATCCGACACAGCTGTCGCGCTGGGTGGCCAC 699
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13527 AACGTCATCCGACGAGGAGGCGGCGGATCCGACACAGCTGTCGCGCTGGGTGGCCAC 13586
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 700 ACGGACTCAGTCGAGGCGGCGGCGGATCCGACACAGCTGTCGCGCTGGGTGGCCAC 759
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13587 TTGGACAGCGCTTCGCGAGGAGGCGGCGGATCCGACACAGCTGTCGCGCTGGGTGGCC 13646
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 760 TTGGTCATTGCCAAGGCGCTACGCACTACTCC---GTCAAGATGCCGCTGGCTTCTC 816
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13647 CTGGAAAGCGCAGTCAGCTGGGGAACCTCAACCGCATGTGTCAACAGCGGTACGCTGCC 13706
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

/Note="PS00758 Arge / dape / ACY1 / CP62 / yscs family
signature 1"
gene complement(4340..6592)
/ gene="pkng"
/ complement(4340..6592)
/ note="Rv0410c, (MTCY22G10.06c), len: 750 aa, protein
kinase similar eg to PKNA_MYCLE P54743 protein kinase
pkna, mycobacterium le (253 aa), fasta scores, opt: 291,
E(): 3.1e-10, (33.2% identity in 203 aa overlap); contains
PS00108 Serine/Threonine protein kinases active-site
signature"
/ codon_start=1
/ transl_table=11
/ product="pkng"
/ protein_id="CAB06580.1"
/ db_xref="GI:1817676"
/ db_xref="SPTREMBL:P96256"
/ translation="MAKASETERSGPGTQPADQATATATVRLSTQAVRPPDGED
NRPPLTGPDPEDQDMATTSRYRPPRYRLGGGLVETPRAPDIPFALMTNPVYES
KRFQWNGRPGVSDSETKASGEGWCPGSPISFQPLNPGDIVAQVEKCIAGS
GLGWITLADNRNNGRPVYKGLVHSGDAEQAMAAERQFLAEVHPSPVQVFNFE
HIDRHGDPVGYIYMEYVGOSLRSKSGOKLPEAALAYLEILIPALSHSISGLVND
LRPENIMLFEOLKIDLGAVSRINSFGYLGPGQAPVIRGPTVADITVGRNT
LAATIDLPTRNGRYVDGLPEDDPVLTQDYSGRILRAIDPDPRQRTTAERMSOL
TGLREVVAADITVPRPGSTSPRSSTFGVDLVAHNDVYIDGQVHAKEKLTANETV
TALSVPVDPDVAASVLDATVLSQVQTLDSLRARHGLDLDGVDSESEVPLME
VRALDLGVAKATRKLDLDERVGRWRLVWRAVAELTGTGDIHETFEVDLTFE
PGEIARLALAAELAGNTEDEHEFYQVSTWSTGDIYSAFGLARARSAGDVGAIVR
TIDVEVPTSRHETARTSATVLSGSRSEVTEQIRDAARVEALPPEPVQIDR
ALVIGALMDIMKONKASTNHILGFPFTHGLRIGVASLSRLARVAPLQRHRTLVDM
ANKVRSPTSF"
misc_feature complement(5741..5779)
/ gene="pkng"
/ note="PS00108 Serine/Threonine protein kinases
active-site signature"
/ complement(6592..7578)
/ gene="glnH"
/ complement(6592..7578)
gene / gene="glnH"
/ note="Rv0411c, (MTCY22G10.07c), len: 328 aa, glnH,
similar to eg GLNH_BACST P27676 glutamine-binding protein
precursor(262 aa), fasta scores, opt: 493, E(): 7.5e-22,
(37.8% identity in 193 aa overlap), contains PS00013
Prokaryotic membrane lipoprotein lipid attachment site,
PS01039 Bacterial extracellular solute-binding proteins,
family 3 signature"
/ codon_start=1
/ transl_table=11
/ product="glnH"
/ protein_id="CAB06581.1"
/ db_xref="GI:1817677"
/ db_xref="SPTREMBL:P96257"
/ translation="MTRRALIARAAPLALAVLAVASGHSSETLVEATPLPPT
PVGMEIMPPOPPLPPDSSODDPTASLRPAKRAADAADNIRARGILVIGLIDGS
NLSFSDPTGTGTVGVDIAGVARDICGVPRVYHRTLSAERTALOKSDVDIVY
KIMSTICERKLVNFTVYLDANQRIARPDSFTTVSDISGRKVCVAAKGTISLRNIR

```

```

QY 817 TTCTGACACGACGAGAGAGTTCGCTGCTGGCGACGCACTACTACTCTCCATCTGAAT 876
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13707 TTCTGGGGCGCCGACGAGATTCGCTGATGCTACAGAACTACTCTGCTGGAGAC 13766
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 877 GCCACCGAGTGAACAACATCCGATCTGACCTGACCTGACATGATGCGCTCACCTAAC 936
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13767 ATGACAGCGCTCAAGAGCATGCGCTGTATCTTAACCTTGACATGTTGGCGTCCGAAAC 13826
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 937 TACGCCCTCATGATCTATGACGAGTGTATGATCGCGC-----TTCAACAG 981
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13827 CCGGATTACTTCACTACGACGAGTGTACAGAGTGTACAGAGTGTACAGAGTGTACAGAG 13886
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 982 ACGGACCGCGCGCTTCGCCGACATCGAAGAACTGTGAGAGACTATACACTCATC 1041
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13887 GTGGTGGCCCAAGAGCTCGCGGATTCGAGCGCAGCTGCTCCATCTGAAGATGGCC 13946
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1042 GACCTGCTCATATCCCAACCCAGTTTGAAGGAGCTTCGAGTACGAGGCGCTTATCTG 1101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13947 GCGAAGACCGCGGACGACACCTGTTGACGCTGCTCGACTACGACGCTTACAGCTG 14006
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1102 AAGGCAATTCGTCGCTGAGTCTTACGCGGCGCGGAGGCGCATCATGTCGAGAGAAC 1161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 14007 GCGGATATCCCTTCGCGGCTGCTTCTCGCGCTGAGGTCAAGAGTCCCGCGAGCAA 14066
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1162 GCAAGCGCTGGGAGGTCAAGCGCGGCTGCTACGAGCCCACTACACGCGCGGGGA 1221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 14067 GCGGAGTCTGCGGCGGCGACCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 14126
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1222 GACACATGACCAACTCAACATGAAAGCCTTCATCACTCACTCACTCACTCACTCACT 1281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 14127 GACACCTGACCATATATGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 14186
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1282 GCGGTGCGCACCTACGCGCAAGAGCTC 1308
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 14187 GCGGTGCGTTGTATGCGCAGAGCTC 14213
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
LOCUS YSCAMY 2272 bp DNA 05-JUL-1994
DEFINITION Saccharomyces cerevisiae aminopeptidase Y gene, complete cds.
ACCESSION U31635
VERSION U31635.1 GI:469463
KEYWORDS aminopeptidase Y.
SOURCE Saccharomyces cerevisiae
ORGANISM Saccharomyces cerevisiae
REFERENCE 1 (bases 1 to 2272)
AUTHORS Nishizawa,N., Yasuhara,T., Nakai,T., Fujiki,Y. and Ohashi,A.
TITLE Molecular cloning of the aminopeptidase Y gene of Saccharomyces
cerevisiae. Sequence analysis and gene disruption of a new
aminopeptidase
JOURNAL J. Biol. Chem. 269 (18), 13651-13655 (1994)
MEDLINE 94230479
FEATURES
source
location/Qualifiers
1..2272
/organism="Saccharomyces cerevisiae"
/strain="D273-10B"
/db_xref="taxon:4932"
/map="Chromosome II"
121..126
220..282
220..1833
/EC="number=3.4.11"
/Note="amino acid feature: potential N-glycosylation site
1, bp 472..474; amino acid feature: potential
N-glycosylation site 2, bp 505..507; amino acid feature:
potential N-glycosylation site 3, bp 562..564; amino
acid feature: potential N-glycosylation site 4, bp 667..
669; amino acid feature: potential N-glycosylation site 5,
bp 703..705; amino acid feature: potential

```

[illegible]

Query Match	11.4%	Score 163.8	DB 7	Length 2272
Best Local Similarity	52.9%	Pred. No. 1.1e-21		
Matches 407; Conservative	0	Mismatches 347	Indels 15	Gaps 2

OY	374	TCGCGCTGATCAAGAGGTGGAGAATGCCGTTGGGGAGCAAGTCGGTCCGTCGAAG	433
Db	833	TTGCCCTTATCTCAAGAGAGTAAAGTGTCTCTTTGGGGATTAAGTAAATTATTTGGCCGGTAACT	892
OY	434	CCAAAGCCGCGGCTTCGATTGTCTATTAACAATGTGGCGGATCCATGGCCGGCACCTTGG	493
Db	893	TCGGTTTCAACCCCTGTGGCTATTTACGACACGAACCTAAATCTAAGAGAGACTACACG	952
OY	494	GGCGGGCGAGAGTGATAAAGGAGCCGATATGGCCATTGTGGTATACAGCTTGAGGAGATG	553
Db	953	GTAACATTAGGCGAACCTTACCAAGCACACTGTGGCCACCGTCCGGTGTTCATATTAAGTTG	1012
OY	554	GCCAGAACTGAT---CAAGCTTGCTGAGCGCTGATCCGGTATCTGTGCACTGTGGGTGG	610
Db	1013	GTAATAAATTTGATTTGCCATATTTGCATTTGACATGATGATATATCATTAATATTTTGGCAATG	1072
OY	611	ATAGTAACGAGAGAAACCGTACGAGCTATTAACGTGTGGCGGACGAAAGGCCGGCGATC	670
Db	1073	ACTCTTATGTAGAAATTTATCAGAGCCCAAGAAATATTTGCCAGCACAAACACGGTGTATC	1132
OY	671	CGAACACAGCTGCTCGCGGTGGGTGGCCACACGAGCTCAAGTCGAGGCGGGCCCTGGTATCA	730
Db	1133	CTGATAAATATTGTTCCCTTGGTGGCTCATTTACAGACTCTGTGTAGGAGGGGCCAGGATATCA	1192
OY	731	ACGAGATGGGCTCGGGGCAATTATTACCACTGGTCTATGTGCCAAACGGCTCACAGCACTCT	790
Db	1193	ACGAGATAGGTTCCGGTACTATTTCCCTATTTAAGCTGCTTAAGCAATTGACCCATTTC	1252
OY	791	CCGTCAAGATGCCCTGGCGCTTCTCTTCTGGACACAGAGAGTTCGGTCTGCTGGGCA	850
Db	1253	AGATCAATAAACAGAGTCCGTTTGGCATGGTGGGCTGCTGAAGAAAGATTGTCTAGAGCT	1312
OY	851	GCAACTACTAGCTCCCATCTGAATGGCACCGGAGCTGAACAAGATCCGAGCTGACTACGA	910
Db	1313	CTAACTCTACGCTTATTAACCTTGACCAAGAAAGAAACCTCAAGATCTCAAGATATTATTAAG	1372
OY	911	ACTTCGACATGATCGCTCTACACTTAACCTACGCGCTCATGATCTATGATACGCGTGTGATCGG	970
Db	1373	ACTATGACATGATGCTCTCTCCAACCTTGAATACGAATTTATGAT-----TG	1420
OY	971	CGTTAACACAGAGCGGACCGCGCGGTTCCGCCCAATGAGAAATGTTTCCAGAGCTACT	1030
Db	1421	CGAACAAACAGAGAACCTTAAGGGGCTGGAAGATTTGAAAAAAGCTGTACGAGACTACT	1480
OY	1031	ACGACTCCATGACACTGCTCATATCCCCACCCAGATTGTGACGAGAGCTTCCGACTCGAGG	1090

[illegible]

RESULT	3
SCYBR286W	
LOCUS	
DEFINITION	S. cerevisiae chromosome II reading frame ORF YBR286w.
ACCESSION	Z38155.Y13134
VERSION	Z38155.1 GI:1870117
KEYWORDS	.

ORGANISM
 Saccharomyces cerevisiae
 Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
 Saccharomycetaceae; Saccharomyces.
 1 (bases 1 to 2532)
 Brandt, T., Christiansen, C., Holmstroem, K. and Kallseoe, T.
 Unpublished
 2 (bases 1 to 2532)
 MIPS.
 AUTHORS

TITLE Direct Submission
JOURNAL Submitted (30-AGO-1994) Data collected by MIPS on behalf of the European yeast chromosome II sequencing project. MIPS at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152 Martinsried, FRG; E-mail: Mewes@mips.embl.de.org
REFERENCE 3 (bases 1 to 2532)
AUTHORS Feldmann, H., Aigle, M., AjlinoVIC, G., Andre, B., Bactet, M. C., Barth, C., Baur, A., Becam, A. M., Bileau, N., Boles, E., Brandt, T., Brendel, M., Brueckner, M., Bussereau, F., Christensen, C.

Cortieras, R., Crouzet, M., Czapluch, C., Demolis, N., Delaveau, T.,
Dognon, F., Dondey, H., Ducrest, S., Dubois, E., Duyon, B., El
Bakoury, M., Entlan, K. D., Feuerann, M., Fiers, W., Fobo, G. M.,
Fitz, C., Gassenhuber, H., Glansdorff, N., Goldau, A., Grivell, L. A.,
de Haan, M., Helm, C., Herbert, C. J., Hollenberg, C. P., Holmstrom, K.,
Jaeq, C., Jacques, M., Jaunault, J.-C., Joniaux, J. L., Kallesoe, T.,
Klesau, P., Kirchbach, L., Koeter, P., Koris, S., Liehl, S., Logny, M.,
Lohar, A. J. E., Louis, E. J., Li, Z. Y., Maat, M. J., Maillet, L.,
Muller, J. M., Noels, E., O'Connell, T.

Nasi, F., Obermayer, B., Perez, J., Pierard, A., Pirivadi, E., Pohl, F.M., Pohl, T.M., Potter, S., Proft, M., Purnelle, B., Ramezani Rad, M., Rieger, M., Rose, M., Schaaf-Geertsenschlager, I., Scherens, B., Schwarzlose, C., Skala, J., Slonimski, P.P., Smits, P.H.M., Souclet, J.L., Steensma, H.Y., Stucke, R., Tarrazarzu, A., van der Aart, O.J., van Dyck, L., Vassarotti, A., Vetter, I., Vierendeels, F., Vissers, S., Wagner, G., de Wegdijosse, P., Wolfer, K.H., Zaigalski, M., Zimmermann, F.K., Mewes, H.W. and Klein, K.

JOURNAL	EMBO J. 13 (24), 5795-5805 (1994)
MEDLINE	95112788
COMMENT	On Mar 5, 1997 this sequence version replaced gi:536735.
FEATURES	Location/Qualifiers
SOURCE	1..2532 /organism="Saccharomyces cerevisiae" /db_xref="taxon:4932" /chromosome="II"

```

CDS
<1..348
/note="ORF YBR285w"
/codon_start=1
/protein_id="CA85250.1"
/db_xref="GI:1870118"
/db_xref="SPTREMBL:P89508"
/translation="SLESPTDDSMEDISHRENATILTRNRNNGKVALDKGVNNHC
IDNEELKGLVANEHKLNDNSTRHONTSPEDSEVDFRDDQKSRILKSTRRSYLR
YK"

```

```

gene 662..2275
      /gene="APE3"
CDS 662..2275
      /gene="APE3"
      /note="ORF YBR286w"

```

```

/codon_start=1
/db_xref="GI:1870119"
/translation="MHFSLKOLAAVAFAATMGSAYVLPQFOEAFQOEPIENYLPQ
LNDSSAVANIRKPHIPIYPMKPHVESEKLODKIKVDLWATMDLYRLANTSTPDY
GHPTRVSGKHNKMTMEYILNVPDMDQDYVYSIQEPALSGKTIISFYLSDAENGKSP
ANTTAPALSPYDGFVGLKVIPIPLGCEKRYASVPRHREKIALIERKCFPGDK
SNLAGKFGTAIVVLYDNEPKSEGLHGLPTHTVATVAPYKVGKLIANTALNI
DYSLYFAMDSEYVEIKTONIADTKHGPDPNIVAGHSDSEVPGINDGSGTISL
LNVAEKIDHFKIKNKVFAMMAAESEGLGNSFYAGNTEKESKIRFYMDYDMMAS
NYXEIYDANKENPKSEELKNLYDYKKAHNLNYLTPVEDGSDYGFNNGIPPG
GIATGAKNNVNNKVLDRCYHOLCDVSNMAFIINTKLIAHSVATYDSREGPR
KRETOKEHEDILNAQDQFERYRDFLIT"
BASE COUNT      829 a      517 c      482 g      704 t
ORIGIN

Query Match      11.4%; Score 163.8; DB 7; Length 2532;
Best Local Similarity 52.9%; Pred. No. 1.1e-21;
Matches 407; Conservative 0; Mismatches 347; Indels 15; Gaps 2;

Db 374 TCGCCCTGATCAAGCGTGGAGATCCCGTTGGCGACAGTCCGTTCTCGCTGCCAAG 433
1275 TTGCCCTATGAAAGAGTAGTGTCTTTGGCGATAGAGTAAATTGGCCGGTAGT 1334
Qy 434 CCAGGCGCGGCTTCGATGCTATTAACATGTGGCGGATCCATGGCGGACACCTTG 493
Db 1335 TCGGTTTACCCCTGTGGTCATTTACGACACGAACTAAATCTAAGAGACTACAG 1394
Qy 494 GCGGCGCGAGATGATAGGACCGTATTTGCCATTGTGGCTATTCAGCTGGAGAGATG 553
Db 1395 GTACATTAGGCGCAACCTACCAAGCACTGTGCCACCGTCCGTTTCATATTAAGTTG 1454
Qy 554 GCCAGAACCTGAT---CAAGCTTGTGAGGCTGATCGGATCTGTGATCTGTGGGTTG 610
Db 1455 GTAAAAATTTGATGCCAATATTTGATGACATGATATTCATTATTTTGGCATNG 1514
Qy 611 ATAGTAGCAGAGAACCGTACGAGCTATACGTTGTGGCGGACAGAGGCGCGGATC 670
Db 1515 ACTCTTAGTAGAATTTATTCAGAGCCAGACATATTCGCCACAAACACGGTGATC 1574
Qy 671 CGAACAGCTGTCGCGGTGGGTGGCGACAGGACTCAGTGGCGGCGGCGCTGGATCA 730
Db 1575 CTGATTAATTTGTCCTTGGTGTCTCATTCAGACTCTGTTAGAGGCGCCAGGATCA 1634
Qy 731 ACGAGATGGCTCGGGCATTTATTAGCAATTGTCATTGCGCAAGCGCTCAGCAGTACT 790
Db 1635 ACGAGAGTGTCCGGTACTATTTCCTATTAAGAGTGGCTAAGCATTTGACCAATTTCA 1694
Qy 791 CCGTAAAGATCCGTCGCTTCTCTTGTGACAGCAGAGAGATTCGCTGCTGGGCA 850
Db 1695 AGATCAATAAACAGTCCGTTTTCATGAGTGGTGTGCAAGAGAGATTTGAGGCT 1754
Qy 851 GCAACTAGTACGTCCTCATCGAATGGCAGGAGCTGAAACAAGATCCGACTGTACCTGA 910
Db 1755 CTAACTTACGCTTATATCTTGACCAAGAAAGAAATCCCAAGATCAGATATTATTA 1814
Qy 911 ACTTGACATATCGCTCACCCTAATCAGCCCTCATGATCTATCTATGACGCTGATGATCG 970
Db 1815 ACTATGACATATGCTTCTCCAACTATGAAATAGAAATTTATGA-----TG 1862
Qy 971 CGTTCAACGAGCGGACCGCGGCTTCGCCAGATCGAGAACTGTTGAGAGACTACT 1030
Db 1863 CGAACAAAGAGAGAACCTTAAGGGGCTGGAAGAGTTGAAAAACCTGTAGCTAGACTACT 1922
Qy 1031 ACGACTCATGAGCTGCTCATATCCCAAGGAGTTGAGGAGAGTTCGAGCTAGAGAG 1090
Db 1923 ACAAGGCTCATGCTGAATCACTTTGGTCCCATTTTACGATGATGATGATGTCG 1982
Qy 1091 CCTTATCTGAAGGCAATTCGTCGGGTGAGACTCTTCACGGGCGCGGA 1139
Db 1983 GGTATTATCAACAAGGAATTCAGCGCGGTGTATTGCGACCTGTGCGCA 2031

RESULT 4
LOCUS SCDBP3 32421 bp DNA PLN 05-DEC-1994
DEFINITION S.cerevisiae (s288c) RFL1, DPB3, YML27 and SNF5 genes.
ACCESSION X76053
VERSION X76053.1 GI:600025
KEYWORDS DPB3 gene; RFL1 gene; SNF5 gene; YML27 gene.
SOURCE Baker's yeast.
ORGANISM Saccharomyces cerevisiae
Eukaryote; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 32420)
AUTHORS Holmstrom,K.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-1993) K. Holmstrom, Biotechnological Institute,
Lundtoftevej 100, Building 227, 2800 Lyngby, DENMARK
REFERENCE 2 (bases 1 to 32420)
AUTHORS Holmstrom,K., Brandt,T. and Kallioe,T.
TITLE The sequence of a 32 420 bp segment located on the right arm of
chromosome II from Saccharomyces cerevisiae
JOURNAL yeast 10, 47-62 (1994)
REFERENCE 3 (bases 1 to 32421)
AUTHORS Holmstrom,K.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-1994) K. Holmstrom, Biotechnological Institute,
Lundtoftevej 100, Building 227, 2800 Lyngby, DENMARK
On Dec 7, 1994 this sequence version replaced gi:429119.
FEATURES
location/Qualifiers
1..32421
1. /organism="Saccharomyces cerevisiae"
/db_xref="s288c"
/db_xref="taxon:4932"
complement(1..3289)
/gene="Rfl1"
/db_xref="SGD:S0000479"
complement(<1..3289)
/gene="Rfl1"
/note="YBR1743-ORF"
/codon_start=1
/db_xref="SGD:S0000479"
/protein_id="CAA53638.1"
/db_xref="GI:429120"
/db_xref="SWISS-PROT:P29539"
/translation="MSKDFSDKRRHTIDRIDHILRSQDNYSNGSSPMKATNLP
SPQAMHIQDLSPTPKRRKLASSSDCENQFDLSANKMLYEDTGRMLQSLPELS
ASNDNAPYTKSYAFSDRIESSPIYIPGSSPPSPSKGKILNRPLRYATVD
LSYKRLQYTHKHLNGNITFSPYKRETVNRALEWVGSEIHLGVNDESEFEIEI
GGGLILOESEDYARREFYATPNNIIPITLTKNVNDOKRNLIVNESEIEICI
PHLDAODTLSSSEKKNPFVIRLVIVVFESKINSNFKIVMLTRPDLVNLKAI
YRWTTGLRNENSKRIITAOVSFLRDEKGTFFLISNEIKPIISTETMEINSHV
IYKTLILRFLSKYPLMTEYVSWLPGVLPILIGDIYSKILITISIVYLELL
KICLDFEDEHERIYOCIMLSPVCEIIEKFLSKPLIISYDSANIDKYLIGHLQOIK
NYIVKNDKNIADMLWISMTGLAYDSGRVYDLTSESKYWFPLNNICFIIINHKTL
MSIKVARIITVCICTKRSOKNOBKSLSLPRPMTQPYVNDPBARGCIIHLG
VVIATFSKRLSDMEELFDHDLITITIEDYIFKDISILQNVLFVYLLIGKRA
DVALERKYKRIHPMSVIASEGVLKLISSLPQIIRKEDKILKVVQVEVAISV
NLARDLILTSILKIPEDRKDOTHLESSESLIKVTONNKDPIERDFGVAISFVY
FLDLFLKNDSSLVFNISQISKVIISOGNNLTDLKQVIRKARETSEFLIEKFL
DDKTEVYAGNWGSTLLPNISPREQSLANIVKYPNSENIEFLDLCKISFPVN
LFTLLHYSMSNNNFYFILOSYSKKNKKNIVDLITLTKSLGPNELFSGILPEL
RKFMDITIEYCIHSNPNLNSIPDLNSDLKLKLPKRSASIFPANKIKFSCSDITVR
WLLKGQLELDLNNFSEIENVLQNASSELEKSEIIRLELLHMANPIEPLFGLNF
CIKNNMAD"
complement(3804..6227)
/note="YBR2013-ORF"
/codon_start=1
/protein_id="CAA53639.1"
/db_xref="GI:429121"
/db_xref="SWISS-PROT:P38148"
/translation="MVEVPSITPGLHDLMLRLHDAEMPECKMPFADISFGPP
DFPHSLAIVASQDANNSALLRNSLEVNDLFGQSKVATSFHREDTCEITDNSGFGY

```


[illegible]

RESULT	6	SLAMIC	3480 bp	DNA	BCT	07-SEP-1998
LOCUS		SLAMIC				
DEFINITION		S. lividans amC gene.				
Qy	335	GCAGGAGGCGGATTACCCATTCGATGCGAGGGCAAGGTCCGCTGATCAAGCGTGGAG				394
Db	1188	GCAGGAGGCGGATTACGAGAGGATGCGCGCTGCGGGGGATTCGCGTCTGACAGCGCGGG				1227
Qy	395	AATGCCGCTTGGCGACAGTCCGTTCTGCTGCCAAAGCCAGAGCGCGCTTGATG				454
Db	1228	GCTCGTGGATGAGACAAACACACCGGGATGGCCGCCACCGCCGGCGGCGCGCTGATTCG				1287
Qy	455	TCTATACAAATGTGGCGGAGATTCATGGCGGGACCCCTTGGCGCGCGCGCAAGATGTAAG				514
Db	1288	TGCGCAACACCGGCGCCCGGGAACTGCACACTCGCTGGCCGACCCCGAGCGCGCCGCA				1347
Qy	515	GACCGTATTCGGCCATTTGCGTATGACTTGGAGGATGGCCAGAAAGCTGATCAAGCTTG				574
Db	1348	TAC-----CCATGCGCGGGGTACCCAGAGAGAGCGGCGCGCTCTCGCGCGCGAGG				1398
Qy	575	CTGAGGCTGGATCGGTATCTGTGGATCTGTGGGTGATAGTAAAGAGAAACCTTACGA				634
Db	1399	CACGGGCGGGA--CGCACGCTCCACTACACCTCGCTCGCTCAACCGACGCTCTACCA				1455
Qy	635	CGTATACGTTGTGCGCCACACAAAGGGCGCGCATCCGAACAGCTCGTGGCGTGGGTG				694
Db	1456	CCGAGAACTGATGCGCGGTGAGGCCCCCGGGGCAACCCGAGACACCGTGTCTCGGGG				1515
Qy	695	GCCACACGGACTAGTCGAGGCGGGCCCTTGATTCACAGCAGATGGCTCGGGCATTTATTA				754
Db	1516	CCCACCTCGACTCCGTGCCGAGGGGCGCGGCTACACACACAGCTGTGCGCGCGCGCG				1575
Qy	755	GCACTTGTGATATTGGCCAAAGCGCTACGACGAGTACCTGTCGAAGATGCGTGGCC---				811
Db	1576	TACTGCTCGAAAGGCGCCCTCGCGGCCGCCAAGGACACCGGACCCGGGCGAGGAAAC				1635
Qy	811	-----TTCCTCTCTGTGACGACGAGAGAGTTGCTGTCTGGGCGACCACTACTACG				862
Db	1636	GCGTGTGTTGCGGCTTCTGTGGGCGCGGAGGAGTTCGGCTCTCGGCTCGGACGCTACG				1695
Qy	863	TCTCCATCTGAAATGGCACGAGACTGAAACAATCCGATCTGATACCTGAACTTCGACATGA				922
Db	1696	TGACTCGTTGACCCAGAGAGACGTGACGCGACCGGCTGTACTTCAACCTGGAGATGA				1755
Qy	923	TGCGCTCACCTTACTACGCGCCCTCATGATCTATGACGCTGATGATGATCGCGCTTCAAC---				979
Db	1756	TGCGAGACCCCACTACGCGTCTTCTACGCTGAGATTCCAAAGGCCACCGACCCCGTCAAG				1815
Qy	979	--CAGACGGAGACGGGCGCGGTCCGCCCAATGAGAAATGTTTCAAGAGATCTACGACT				1036
Db	1816	GACGACGCGCGCGCGCGCGGTCTCGAGAAATGAGGCGGAGGTGACGACACGCTTCGCG				1875
Qy	1037	CCATGACACTGCTCATATATCCCAACCACTTTGACGAGAGTCCGACTACGAGGCGCTTTA				1096
Db	1876	CCCAAGGGCGGACCTCCCTCCGCGGAGACCGCGGACGGGCGCTCGACATCACTGAGGTTC				1935
Qy	1097	TCTGAAAGGCATTCGTCGCGGTGGACTTTACGCGGCGCGAGGGCATCATATGCCAAG				1156
Db	1936	TGGCGCGGGCATTCGCCACCGGCGGCTCTTACGAGAGCTGTTCAGGCGCCAAGACCCCG				1995
Qy	1157	AGAAAGCAACCGCTGGGAGGTCACAGCGGGGTGGCCCTACGAGCCAACTACACGCGCG				1216
Db	1996	AGCAGCGCGCTCTGTGTGGGGGAGACCGCTGCGCGCGCGTACGACCCCTCTGTACACCTGA				2055
Qy	1217	CGGAGAGCAACATGACCAACTTCACATGAAGCCTTCTGATCAACTCCAAAGGACCG				1276
Db	2056	CGTCCGACAGAGAGGGCGAGTACAGCCGGAGGCGGACGCTGCACGAGGCGGTTC				2115
Qy	1277	CTTTCGCGCTGCGCACCTACGCCACGAC				1306
Db	2116	AGACGCTGCTACCCACTACTCCCGGACCC				2145

FEATURES	source
gene	1. .3480 /organism="Streptomyces lividans" /db_xref="taxon:1916" complement(1. .26) /gene="am1B" complement(<1. .26) /gene="am1B" /note="EMBL Acc. No. 285949" /codon_start=1 /transl_table=1 /protein_id="CAB06815.1" /db_xref="GI:1881522" /db_xref="SPTREMBL:P97179" /translation="MHGNTSPAR" 407. .3388 /gene="am1C" 407. .3388
gene	407. .3388 /gene="am1C" /codon_start=1 /transl_table=1 /protein_id="CAB06816.1" /db_xref="GI:2780210" /db_xref="SPTREMBL:P96992" /translation="MRMLPGRELTVAAVNAVLASVAPPAPASPADPAAAAAEBGNTA TVFYTTTKMNDRYNLHYAPDGGSGVAGPAAACIDWYKRPGLASAGLATFENFAGVLA NGSTPTNNNGENYALIGALTGVKDVGVAAHSDPDAGODDPTPGDGRARAVVYSTRE LGMTANHHRYPAGSGWTAYPGVGMRAACAKAMKDKVDLGAHSLTAFFNNNGEMNDN NKGDITVLPAGVTIVADTVTADDDVDAEESVPTAFTVPAVSARADGVSVLVLM PSTDTGVAKYQVTRSGGTGEGVSVDSIVSGTGLEERTYTYTVAATDAAGNSVA ASAAAIATGTRGRPPAPAGRLATDPRKDPYFVLLARFNDATNNNGSGHSGSGG AANDPMEGRDGEKGLVNLKLDYIKGIGFSAWITPVILNRSDYDHGXGDFEYVDEP LESAGASLODLIDAAHAKGMKIVODVYVNHSSRMGARGLFTPYVGVADDSQMSWYVE KQPEFEDGLTEPFRKSGSKSYTNGDLNSTAEPTGTCYVMGRPTAKSFEETGLYDCOC PSPISGMEPKALYHNCWTGNWGEDSRSCWLEHDLADNTENPOVQNTLIGATYKIID MGVGFERVDYAVHI.PRTTWNRRFLPAIDERYAQOHGAEEAANFEVGEVAFVNDKNNS RGSTVYAKQFYTWKEREKYESADDKAKAILEMEYEQOOGTGQNPISANFLGNSYJHN DHSFSGKMNVIDMHHMNFEGDASNAFHNGKSDSDSYNDATYVNYVYVSDHGXGDFEYVDEP PESGSGMAEMNMSIMTFERRGIIPTLYLNGSELEPQKRGKIDGSPCLATGRAYGGB LAGVYTSADSRVDSATGAVATTIAGQPYKRVQNLNRRRVYPLAMOGYSTDVTGTC MAFRRYTIDGDTSFALYVSGAAYDGIPINGTYRADVATGDTFRVTDGRLTVAAPGAKK NLRYVYNGPKKIDADGEYLYK"
gene	407. .3388 /gene="am1C" /codon_start=1 /transl_table=1 /protein_id="CAB06816.1" /db_xref="GI:2780210" /db_xref="SPTREMBL:P96992" /translation="MRMLPGRELTVAAVNAVLASVAPPAPASPADPAAAAAEBGNTA TVFYTTTKMNDRYNLHYAPDGGSGVAGPAAACIDWYKRPGLASAGLATFENFAGVLA NGSTPTNNNGENYALIGALTGVKDVGVAAHSDPDAGODDPTPGDGRARAVVYSTRE LGMTANHHRYPAGSGWTAYPGVGMRAACAKAMKDKVDLGAHSLTAFFNNNGEMNDN NKGDITVLPAGVTIVADTVTADDDVDAEESVPTAFTVPAVSARADGVSVLVLM PSTDTGVAKYQVTRSGGTGEGVSVDSIVSGTGLEERTYTYTVAATDAAGNSVA ASAAAIATGTRGRPPAPAGRLATDPRKDPYFVLLARFNDATNNNGSGHSGSGG AANDPMEGRDGEKGLVNLKLDYIKGIGFSAWITPVILNRSDYDHGXGDFEYVDEP LESAGASLODLIDAAHAKGMKIVODVYVNHSSRMGARGLFTPYVGVADDSQMSWYVE KQPEFEDGLTEPFRKSGSKSYTNGDLNSTAEPTGTCYVMGRPTAKSFEETGLYDCOC PSPISGMEPKALYHNCWTGNWGEDSRSCWLEHDLADNTENPOVQNTLIGATYKIID MGVGFERVDYAVHI.PRTTWNRRFLPAIDERYAQOHGAEEAANFEVGEVAFVNDKNNS RGSTVYAKQFYTWKEREKYESADDKAKAILEMEYEQOOGTGQNPISANFLGNSYJHN DHSFSGKMNVIDMHHMNFEGDASNAFHNGKSDSDSYNDATYVNYVYVSDHGXGDFEYVDEP PESGSGMAEMNMSIMTFERRGIIPTLYLNGSELEPQKRGKIDGSPCLATGRAYGGB LAGVYTSADSRVDSATGAVATTIAGQPYKRVQNLNRRRVYPLAMOGYSTDVTGTC MAFRRYTIDGDTSFALYVSGAAYDGIPINGTYRADVATGDTFRVTDGRLTVAAPGAKK NLRYVYNGPKKIDADGEYLYK"
gene	407. .3388 /gene="am1C" /codon_start=1 /transl_table=1 /protein_id="CAB06816.1" /db_xref="GI:2780210" /db_xref="SPTREMBL:P96992" /translation="MRMLPGRELTVAAVNAVLASVAPPAPASPADPAAAAAEBGNTA TVFYTTTKMNDRYNLHYAPDGGSGVAGPAAACIDWYKRPGLASAGLATFENFAGVLA NGSTPTNNNGENYALIGALTGVKDVGVAAHSDPDAGODDPTPGDGRARAVVYSTRE LGMTANHHRYPAGSGWTAYPGVGMRAACAKAMKDKVDLGAHSLTAFFNNNGEMNDN NKGDITVLPAGVTIVADTVTADDDVDAEESVPTAFTVPAVSARADGVSVLVLM PSTDTGVAKYQVTRSGGTGEGVSVDSIVSGTGLEERTYTYTVAATDAAGNSVA ASAAAIATGTRGRPPAPAGRLATDPRKDPYFVLLARFNDATNNNGSGHSGSGG AANDPMEGRDGEKGLVNLKLDYIKGIGFSAWITPVILNRSDYDHGXGDFEYVDEP LESAGASLODLIDAAHAKGMKIVODVYVNHSSRMGARGLFTPYVGVADDSQMSWYVE KQPEFEDGLTEPFRKSGSKSYTNGDLNSTAEPTGTCYVMGRPTAKSFEETGLYDCOC PSPISGMEPKALYHNCWTGNWGEDSRSCWLEHDLADNTENPOVQNTLIGATYKIID MGVGFERVDYAVHI.PRTTWNRRFLPAIDERYAQOHGAEEAANFEVGEVAFVNDKNNS RGSTVYAKQFYTWKEREKYESADDKAKAILEMEYEQOOGTGQNPISANFLGNSYJHN DHSFSGKMNVIDMHHMNFEGDASNAFHNGKSDSDSYNDATYVNYVYVSDHGXGDFEYVDEP PESGSGMAEMNMSIMTFERRGIIPTLYLNGSELEPQKRGKIDGSPCLATGRAYGGB LAGVYTSADSRVDSATGAVATTIAGQPYKRVQNLNRRRVYPLAMOGYSTDVTGTC MAFRRYTIDGDTSFALYVSGAAYDGIPINGTYRADVATGDTFRVTDGRLTVAAPGAKK NLRYVYNGPKKIDADGEYLYK"
gene	407. .3388 /gene="am1C" /codon_start=1 /transl_table=1 /protein_id="CAB06816.1" /db_xref="GI:2780210" /db_xref="SPTREMBL:P96992" /translation="MRMLPGRELTVAAVNAVLASVAPPAPASPADPAAAAAEBGNTA TVFYTTTKMNDRYNLHYAPDGGSGVAGPAAACIDWYKRPGLASAGLATFENFAGVLA NGSTPTNNNGENYALIGALTGVKDVGVAAHSDPDAGODDPTPGDGRARAVVYSTRE LGMTANHHRYPAGSGWTAYPGVGMRAACAKAMKDKVDLGAHSLTAFFNNNGEMNDN NKGDITVLPAGVTIVADTVTADDDVDAEESVPTAFTVPAVSARADGVSVLVLM PSTDTGVAKYQVTRSGGTGEGVSVDSIVSGTGLEERTYTYTVAATDAAGNSVA ASAAAIATGTRGRPPAPAGRLATDPRKDPYFVLLARFNDATNNNGSGHSGSGG AANDPMEGRDGEKGLVNLKLDYIKGIGFSAWITPVILNRSDYDHGXGDFEYVDEP LESAGASLODLIDAAHAKGMKIVODVYVNHSSRMGARGLFTPYVGVADDSQMSWYVE KQPEFEDGLTEPFRKSGSKSYTNGDLNSTAEPTGTCYVMGRPTAKSFEETGLYDCOC PSPISGMEPKALYHNCWTGNWGEDSRSCWLEHDLADNTENPOVQNTLIGATYKIID MGVGFERVDYAVHI.PRTTWNRRFLPAIDERYAQOHGAEEAANFEVGEVAFVNDKNNS RGSTVYAKQFYTWKEREKYESADDKAKAILEMEYEQOOGTGQNPISANFLGNSYJHN DHSFSGKMNVIDMHHMNFEGDASNAFHNGKSDSDSYNDATYVNYVYVSDHGXGDFEYVDEP PESGSGMAEMNMSIMTFERRGIIPTLYLNGSELEPQKRGKIDGSPCLATGRAYGGB LAGVYTSADSRVDSATGAVATTIAGQPYKRVQNLNRRRVYPLAMOGYSTDVTGTC MAFRRYTIDGDTSFALYVSGAAYDGIPINGTYRADVATGDTFRVTDGRLTVAAPGAKK NLRYVYNGPKKIDADGEYLYK"
gene	407. .3388 /gene="am1C" /codon_start=1 /transl_table=1 /protein_id="CAB06816.1" /db_xref="GI:2780210" /db_xref="SPTREMBL:P96992" /translation="MRMLPGRELTVAAVNAVLASVAPPAPASPADPAAAAAEBGNTA TVFYTTTKMNDRYNLHYAPDGGSGVAGPAAACIDWYKRPGLASAGLATFENFAGVLA NGSTPTNNNGENYALIGALTGVKDVGVAAHSDPDAGODDPTPGDGRARAVVYSTRE LGMTANHHRYPAGSGWTAYPGVGMRAACAKAMKDKVDLGAHSLTAFFNNNGEMNDN NKGDITVLPAGVTIVADTVTADDDVDAEESVPTAFTVPAVSARADGVSVLVLM PSTDTGVAKYQVTRSGGTGEGVSVDSIVSGTGLEERTYTYTVAATDAAGNSVA ASAAAIATGTRGRPPAPAGRLATDPRKDPYFVLLARFNDATNNNGSGHSGSGG AANDPMEGRDGEKGLVNLKLDYIKGIGFSAWITPVILNRSDYDHGXGDFEYVDEP LESAGASLODLIDAAHAKGMKIVODVYVNHSSRMGARGLFTPYVGVADDSQMSWYVE KQPEFEDGLTEPFRKSGSKSYTNGDLNSTAEPTGTCYVMGRPTAKSFEETGLYDCOC PSPISGMEPKALYHNCWTGNWGEDSRSCWLEHDLADNTENPOVQNTLIGATYKIID MGVGFERVDYAVHI.PRTTWNRRFLPAIDERYAQOHGAEEAANFEVGEVAFVNDKNNS RGSTVYAKQFYTWKEREKYESADDKAKAILEMEYEQOOGTGQNPISANFLGNSYJHN DHSFSGKMNVIDMHHMNFEGDASNAFHNGKSDSDSYNDATYVNYVYVSDHGXGDFEYVDEP PESGSGMAEMNMSIMTFERRGIIPTLYLNGSELEPQKRGKIDGSPCLATGRAYGGB LAGVYTSADSRVDSATGAVATTIAGQPYKRVQNLNRRRVYPLAMOGYSTDVTGTC MAFRRYTIDGDTSFALYVSGAAYDGIPINGTYRADVATGDTFRVTDGRLTVAAPGAKK NLRYVYNGPKKIDADGEYLYK"
gene	407. .3388 /gene="am1C" /codon_start=1 /transl_table=1 /protein_id="CAB06816.1" /db_xref="GI:2780210" /db_xref="SPTREMBL:P96992" /translation="MRMLPGRELTVAAVNAVLASVAPPAPASPADPAAAAAEBGNTA TVFYTTTKMNDRYNLHYAPDGGSGVAGPAAACIDWYKRPGLASAGLATFENFAGVLA NGSTPTNNNGENYALIGALTGVKDVGVAAHSDPDAGODDPTPGDGRARAVVYSTRE LGMTANHHRYPAGSGWTAYPGVGMRAACAKAMKDKVDLGAHSLTAFFNNNGEMNDN NKGDITVLPAGVTIVADTVTADDDVDAEESVPTAFTVPAVSARADGVSVLVLM PSTDTGVAKYQVTRSGGTGEGVSVDSIVSGTGLEERTYTYTVAATDAAGNSVA ASAAAIATGTRGRPPAPAGRLATDPRKDPYFVLLARFNDATNNNGSGHSGSGG AANDPMEGRDGEKGLVNLKLD

[illegible]

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS.

Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the Frameplot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.gov.jp/>

jun/cgi-bin/frameplot.pl. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid IC2 lies between 809 and 741 on the AseI-B genomic restriction fragment.

Location/Qualifiers

FEATURES

source

1. 42210

/organism="Streptomyces coelicolor A3(2)"

/strain="A3(2)"

/db_xref="taxon:100226"

/clone="cosmid IC2"

1. 1228

/gene="SC1C2.01"

<1. 1228

/gene="SC1C2.01"

/note="SC1C2.01, probable delta-1-pyrroline-5-carboxylate dehydrogenase, partial CDS, len >408 aa; similar to many

e.g. TR:050443 (EMBL:AL010186)

delta-1-pyrroline-5-carboxylate dehydrogenase (M.

tuberculosis) (543 aa), fasta scores: opt: 1753 z-score:

2214.2 E(): 0, 65.1% identity in 407 aa overlap, and

PUT2_HUMAN delta-1-pyrroline-5-carboxylate dehydrogenase

(563 aa), fasta scores: opt: 1273 z-score: 1730.4 E(): 0,

48.5% identity in 408 aa overlap. Contains PS00687 and

PS00070 Aldehyde dehydrogenases glutamic acid active site

and cysteine active site and Pfam match to entry PF00171

aldehyd, Aldehyde dehydrogenases, score 114.70, E-value

1.8e-30"

/codon_start=2

/transl_table=11

/product="delta-1-pyrroline-5-carboxylate dehydrogenase"

/protein_id="CAI19968.1"

/db_xref="GI:3355668"

/db_xref="SPRMBL:086502"

/translation="IDSPCELDIFRFRNVHVARNTLAEPANSGVNNRMDHRLLEG

FVYAITPENFSAIANLPTPALMGNNVWKPSPYQTHAAVILMLQLEAGLPKVIN

LVTGGIAEVALEHRDLAIFHTGSKTFQGLMKTVGNIEKRTIPRLVGEVGG

DELVAHPSADRAVLATLRGAFEYOGKCSASRAIIPASIMNDGFEKFEAAEDYL

TMGVDLDSNFIGAVIDERSFAKNRAKIDRAKEDECITVAGSGYDSDGVEPRTV

ECTDENEYFEFEYGPFLAVHYVDSDADDAAMLTOGESYSDALIGSVISNRAA

AAATMEKRLRYAGNFYINDKSTGAVYGGQPFGGGASGASGINDKAGAPQMLMRTLT

RAI KETLVAPDYTPPHMG"

1. 3807

/note="true overlap with cosmid 8D9"

2. 1189

/gene="SC1C2.01"

/note="Pfam match to entry PF00171 aldehyd, Aldehyde

dehydrogenases, score 114.70, E-value 1.8e-30"

464. 487

/gene="SC1C2.01"

/note="PS00687 Aldehyde dehydrogenases glutamic acid

active site"

548. 583

/gene="SC1C2.01"

/note="PS00070 Aldehyde dehydrogenases cysteine active

stem_loop

1246. 1283

/note="hairpin_loop with 16bp stem"

RBS

1522. 1527

/note="possible RBS upstream of SC1C2.02"

gene

1537. 1668

/gene="SC1C2.02"

1537. 1668

/gene="SC1C2.02"

/note="SC1C2.02, questionable ORF, len: 43 aa"

/codon_start=1

/transl_table=11

/product="hypothetical protein SC1C2.02"

/protein_id="CAI19969.1"

/db_xref="GI:3355669"

/db_xref="SPRMBL:086503"

/translation="MAETVRRVRHLSRTSDSRKNAALQSRSLDRDNGATGH"

1744. 2787

/gene="leuB"

1744. 2787

/gene="leuB"

/note="SC1C2.03, leuB, probable 3-isopropylmalate dehydrogenase, len: 347 aa; similar to e.g. LE03_CORGL

3-isopropylmalate dehydrogenase (EC 1.1.1.85) (340 aa),

fasta scores: opt: 1450 z-score: 1318.1 E(): 0, 68.0%

identity in 338 aa overlap. Contains PS00470 Isocitrate

and isopropylmalate dehydrogenases signature and Pfam

match to entry PF00180 isoch, Isocitrate and

isopropylmalate dehydrogenases, score 336.60, E-value

1.3e-105"

/codon_start=1

/transl_table=11

/product="3-isopropylmalate dehydrogenase"

/protein_id="CAI19970.1"

/db_xref="GI:335670"

/db_xref="SPRMBL:086504"

/translation="MSRSLNLAIVPGDGIQGEVVAEGLKYLAVLPQVVKLEKPEDE

GARRHATGELTLDALDKAHDAIILGAGDSVSGVLERFLLKLRAPDHVN

LRPEKLPGVATPAGIODEITDVVNRGTSPTGNGGTLRKGEHVAEVSNTLY

GPERVVDAPFAPBAAPBKRTLVKKNKNTFNFKYAAEYPRVTDYLV

DATITFLVTPARFDVITDMLFEDITTDLAAVSGSIGAASNINPSGDFSMEEP

VHGSPADGAGKADPTATVSLVALLRLHGLGEDEARIDVADSADIGERGDLPARST

SEIDTAAARVAG"

1759. 2766

/gene="leuB"

1759. 2766

/note="Pfam match to entry PF00180 isoch, Isocitrate and

isopropylmalate dehydrogenases, score 336.60, E-value

1.3e-105"

/gene="leuB"

2458. 2517

/gene="leuB"

/note="PS00470 Isocitrate and isopropylmalate

dehydrogenases signature"

3038. 3042

/note="possible RBS upstream of 11ve"

3052. 4140

/gene="11ve"

3052. 4140

/gene="11ve"

/note="11ve"

/note="SC1C2.04, 11ve, probable branched-chain amino acid

amino transferase, len: 362 aa; similar to many, e.g.

ILVE_BACSV putative branched-chain amino acid

amino transferase (362 aa), fasta scores: opt: 1047

z-score: 1401.5 E(): 0, 44.7% identity in 360 aa overlap.

Contains PS00770 Amino transferases class-IV signature"

/codon_start=1

/transl_table=11

/product="branched-chain amino acid aminotransferase"

/protein_id="CAI19971.1"

/db_xref="GI:335671"

/db_xref="SPRMBL:086505"

/translation="MTPTTELEKRSAPLSDSEAAAILANPGRHRTDMYITKTE

GRGHNDGLVAPYGLSLDPATVWLVHQAQLEIFEGIKAVRDPGSAVTRPKNARQA

SSRRLGMPLELPVDTFIEACDALVAODEKWPVARGESLYLRPFMATIEVGLGVRPAN

misc_feature

misc_feature

misc_feature

misc_feature

CDS

gene

RESULT	8				
LOCUS	SCF7	16911 bp	DNA	BCT	10-MAY-1999
DEFINITION	Streptomyces coelicolor cosmid E7.				
ACCESSION	AF049819				
VERSION	AF049819.1	GI:4803679			
KEYWORDS	ABC transporter; AraC-family transcriptional regulator; ATP-binding; dihydroxy acid dehydratase; DNA-binding; glycine-rich secreted protein; hydrolase; llyd; integral membrane; isochlorismatase; proc; proline-rich; pyrroline-5-carboxylate reductase; repeat region; serine/threonine protein kinase; tetr family; ttps; tryptophanyl tRNA synthetase.				
SOURCE	Streptomyces coelicolor A3(2).				
ORGANISM	Streptomyces coelicolor A3(2)				
REFERENCE	Actinomycetales: Streptomycineae; Streptomycetaceae; Streptomyces 1 (Phases 1 to 16911)				
AUTHORS	Redenbach,M., Kieseir,H.M., Denapalte,D., Eichner,A., Cullum,J., Kinasch,H. and Hopwood,D.A.				
TITLE	A set of ordered cosmids and a detailed genetic and physical map				

```
/gene="SCET_01", trtps, tryptophanyl tRNA synthetase, len
/notes="SCET_01, trtps, tryptophanyl tRNA synthetase, len
>208aa; similar to many eg. SW:SIW_BACST trtps,
tryptophanyl tRNA synthetase from Bacillus
stearothermophilus (328 aa) fasta scores; opt: 576,
z-score: 685.7, E(): 7.2e-31, (47.8% identity in 203 aa
overlap)."
/codon_start=2
/transl_table=11
/label=trtps
/product="tryptophanyl tRNA synthetase"
/protein_id="CAB42660.1"
/db_xref="GI:4803680"
```

gene
CDS

/translating="TRYPRIMADIIAYCTDEPVEDTOVEVLARDAIYAFNORVGRH
TIVVPRAPIPAVAARYMINIQEPASAKMGSSDDTGPPVTILBEPDYVRKKRYATVDGSE
ROVVYDVPERAGLANLLETLAICTDGEPEALGGVDSYGALKRDIKAEVYEMLPVRBS
RMHEISADPGYDVCGYLREGAEKRAMARPTVDADARAIGLLPVPVAAAR"
complement(583 . 1587)
/gene="SCE7_02c"
complement(583 . 1587)

misc-feature
CDS

/note="SCE7_02c, probable Arac-family transcriptional
regulator, len: 334aa; similar to many eggs. TR:CAB40696
(EMBL:AL049587) putative transcriptional regulator (335
aa) fasta scores: opt: 643, z-score: 750.6, E(): 0, (39.9%
identity in 316 aa overlap) and TR:O30507 (EMBL:AF008308)
from Pseudomonas aeruginosa (328 aa) fasta scores, opt:
360, z-score: 423.1, E(): 3e-16, (27.1% identity in 328 aa
overlap). Contains Pfam match to entry PF00165 HTH_2,
Bacterial regulatory helix-turn-helix proteins, arac
family. Also contains probable helix-turn-helix motif
(+4.06 SD) at 241-262aa."
/codon_start=1
/transl_table=11
/label=SCE7_02c
/product="putative Arac-family transcriptional regulator"
/protein_id="CAB42661.1"
/db_xref="GI:4803681"

gene
CDS

/translating="MAGAASAAPRYRALVALFPCIRAFDVAVTEWGTTRTRGAR
AFDRVNAADGTTAVPMRGSLALHPDRPLDWLAEDLVVVGDLDDHLTPAPAFLAL
RAHRGPITVAALICGAFTLAQGLDERRALTHTRLDLRTRPRTVDPDLFTIE
DNINWTACTAGIDLCILTVLAGIAEAAATIASNTAPFPQTQOFENPHHA
DBDADLAGEVBHALRHLDPELTADVADLARAGMSRSRARFFAATGTPPLMLDOR
IAAQKLERTLDLPMPEVARARFGSGEVMMQHFAARLATSPRAYRAFVGGSSPIAR
"
complement(613 . 873)
/gene="SCE7_02c"
/note="Pfam match to entry PF00165 HTH_2, Bacterial
regulatory helix-turn-helix proteins, arac family, score
7620, E-value 6.7e-19."
1662 . 2243
/gene="SCE7_03"
1662 . 2243
/gene="SCE7_03"
/note="SCE7_03, possible hydrolase, len: 193aa; similar to
many eggs. TR:O07081 (EMBL:U03876) hypothetical protein
from Bacillus subtilis (187 aa) fasta scores: opt: 370,
z-score: 455.2, E(): 5e-18, (35.1% identity in 188 aa
overlap) and SW:EMB1_ECOLI Isochorismatase from
Escherichia coli (285 aa) fasta scores: opt: 133, z-score:
168.5, E(): 0.046, (29.1% identity in 196 aa overlap).
Contains Pfam match to entry PF00857 Isochorismatase,
isochorismatase family."
/codon_start=1
/transl_table=11
/label=SCE7_03
/product="putative hydrolase"
/protein_id="CAB42662.1"
/db_xref="GI:4803682"

gene
CDS

/translating="MDIAENMALVVYDVQKGFEEDVFNGTRNPADAINDANQ
SAGRPFVFRHDSDVOAGSLREGEYGNFKFEVERDRKGGAEILTYKNVSFLFAS
POLGAMLRASVLIOTNCMEVETTRARGULGYEVVYPIDATYTFGLEGFCGSC
ROSADIELAASAVSLHGGSFAHYVTTEVIAGS"
1683 . 2153
/gene="SCE7_03"
/note="Pfam match to entry PF00857 Isochorismatase,
isochorismatase family, score 50.00, E-value 5.2e-11."
complement(2307 . 3161)
/gene="SCE7_04c"
complement(2307 . 3161)
/note="SCE7_04c, proc, pyrroline-5-carboxylate reductase,
len: 284aa; highly similar to many from prokaryotes and
eukaryotes egs. SW:PROC_CORBL proc,
pyrroline-5-carboxylate reductase from Corynebacterium
glutamicum (270 aa) fasta scores: opt: 735, z-score:

Query Match	Best Local Similarity	Matches	Score	DB	Length
71 AACGCTGAGAGCTGCGCTATGCTACCCCGAGCGCAATCGCGTCTTGTTGGTAAAG	45.0%	197; Conservative	3.68; 52.4; 0.96; 241; 0; 0;	DB 1; 16911;	
1046 ACCAGGTGACGAGGAGGTGCTATACCGCGCGGTGCGGCGGCTCCAGATTTGTCG					1105
131 CCCACGACGACCGTTAACTATCTTACGAGAGCTGAAGAGACTGGCTACTATGATG					190
1106 TCTCTGATGAAGAGGGGCTCCGGGACGACCTGACCCGGGGGTGACGGGTCCGACAGAG					1165
191 TCTTACAGCACTTCAGGTGACCTGTGTGAGCAATGCCGACCAACGCTCAAGTGGCG					250
1166 TCGAGCAGGCGCCAGTGGTATGCGCGGCGCGCTGACACACCCGCGCTGGCGAGG					1225
251 ATGAGGAATGAGAGCGAGACATGACCTTACCTCCAGCGTGGAGGTACCCCGATG					310
1226 GTGAAGCGGCTCCGACAGAGGCGCGACAGGTGTGCTCCGCGGTGTGGCGCGAGG					1255
311 TAGCCGTGCTCAAGACCTGAGATGACGAGCGGAGATTACCTCCGATGTCGAGGCA					370
1286 GCGGCGAGGACCGGGGGGGGGGGGTGAGTGTGATGATGAGAGCGCGGGGACGACG					1345
371 AGGTGCGCCCTATCAAGCGTGTGAATATGCCGTTCCGCGACAACTCGTTCCGTGCA					430
1346 AGGTTCGGCTTCCGAGACCAATCGAGGGAGCGTGGGGTGTGAAGGCGAGGCGCGC					1405
431 AAGCGAAGCGCGCGCTTCAATTGTCTAATAACATGTGGCGGATTCATGGCGGGAACC					490
1406 ATCGGAGAGCGGTGTGTGCTCGCGCGCGACGCGGCGCAGGTGGAAGCGGGCGCGG					1465
491 TTGGCGGCGCGCAGATG					508
1466 CGGTGCGGTGCGGTG					1483

Query Match 3.48: Score 49.2; DB 7; Length 3932;

Best Local Similarity 50.98; Pred. No. 4.4; Mismatches 113; Indels 0; Gaps 0;

Matches 117; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 1134 CGCGAGGAGGATCATGTGTCCGAAGAGACGCAAGCGGTGGAGAGTCAAGCGGCGTGGC 1193

DB 1819 CTTCATGCCCATCAGCACCAGCTTAAGCTCAACCAAGCAGTGCACACAGCATGCC 1878

QY 1194 CTACAGCCCACTACACAGCGCGGGGAGACACATGACCACTCAACATGAAGCCTT 1253

DB 1879 CACCAACACCTCTTAACCTCAACCGAAGCAGCAGCAGCCAGCAGCAAGCAGCGC 1938

QY 1254 CCTGATCAACGCAAGCAGCGCGTGGCGTGGCGTCAAGCGCAGCAGCTCTCTC 1313

DB 1939 CACCAACGAGCAGCAAGCAGCTCAATGTCCATCAGCAGCTCAAGCAGCAGCAGCTC 1998

QY 1314 GATCCCAAGCAAGATACACATCTCTTCACCGAGCAGCAGCAGCAGCA 1363

DB 1999 CAAGCCTGTCCACACATGTCTCAGCTCCAAACCTGCTCCAGCAGCA 2048

RESULT 11

LOCUS SCH69 35824 bp DNA BCT 15-JUN-1999

DEFINITION Streptomyces coelicolor cosmid H69.

ACCESSION AL079308

VERSION AL079308.1 GI:5102782

KEYWORDS 3-Hydroxyacyl-CoA dehydrogenase; bkdA2; dioxxygenase transcriptional regulator; El-alpha branched-chain alpha keto acid dehydrogenase; PtsW/BodA/SpoVE; glutamine amidotransferase; integral membrane isomerase; penicillin-binding protein; peptidyl-prolyl cis-trans isomerase; protein phosphatase; rRNA methylase; serine/threonine protein kinase.

SOURCE Streptomyces coelicolor A3(2).

ORGANISM Streptomyces coelicolor A3(2)

REFERENCE 1 (bases 1 to 35824) Actinobacteria; Actinobacteriales; Streptomycetaceae; Streptomyces.

REFERENCE 2 (bases 1 to 35824) Redenbach, M., Kieser, H.M., Denapalte, D., Eichner, A., Cullum, J., Kinashi, H. and Hopwood, D.A.

REFERENCE 3 (bases 1 to 35824) A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome

REFERENCE 4 (bases 1 to 35824) Mol. Microbiol. 21 (1), 77-96 (1996)

REFERENCE 5 (bases 1 to 35824) 97000351

REFERENCE 6 (bases 1 to 35824) Bentley, S.D., Parkhill, J., Barrell, B.G. and Randsdram, M.A.

REFERENCE 7 (bases 1 to 35824) Direct Submission

REFERENCE 8 (bases 1 to 35824) Submitted (15-JUN-1999) Streptomyces coelicolor sequencing project, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof. David A. Hopwood, [3] John Innes Centre, Norwich Research Park, Colney, Norwich, Norfolk NR4 7UH, UK

REFERENCE 9 (bases 1 to 35824) Notes:

REFERENCE 10 (bases 1 to 35824) Streptomyces coelicolor sequencing at The Sanger Centre is funded by the BBSRC and Beowulf Genomics

REFERENCE 11 (bases 1 to 35824) Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web.

REFERENCE 12 (bases 1 to 35824) (URL: <http://www.sanger.ac.uk/Projects/S.coelicolor/>)

REFERENCE 13 (bases 1 to 35824) CDS are numbered using the following systems eg SC787.01c, SC.15. coelicolor, 787 (cosmid name), .01 (first CDS), C (complementary strand).

REFERENCE 14 (bases 1 to 35824) The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous.

REFERENCE 15 (bases 1 to 35824) The length in codons is given for each CDS.

REFERENCE 16 (bases 1 to 35824) Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database.

REFERENCE 17 (bases 1 to 35824) The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon.

FEATURES

source

1..35824 /organism="Streptomyces coelicolor A3(2)"

1..35824 /strain="A3(2)"

1..35824 /db_xref="taxon:100226"

1..35824 /clone="cosmid H69"

1..35824 /gene="bkdA2"

1..35824 /complement(1..784)

1..35824 /note="SCH69.01c, bkdA2, El-alpha branched-chain alpha keto acid dehydrogenase, len=261aa; very similar to the first 261aa of TR:Q53592 (EMBL:U017169) bkdA, El-alpha branched-chain alpha keto acid dehydrogenase from Streptomyces avermitilis (381 aa) (fasta scores: opt: 1602, z-score: 1786.0, E(): 0, (90.8% identity in 261 aa overlap). Contains Pfam match to entry PF00676 El.dehydrog. Dehydrogenase El component."

1..35824 /codon_start=1

1..35824 /transl_table=11

1..35824 /product="El-alpha branched-chain alpha keto acid dehydrogenase"

1..35824 /protein_id="CAB45198.1"

1..35824 /db_xref="GI:5102783"

1..35824 /translation="MTVMKQKAYRPTPPAPQPTDPAIPDTPPPRIVCTKADKADTLELRTLYALVGRYRNYAQTALTKGRLAVPSSGQACIAALAEEDWLFPSTRTLVVSGVDPVYALTLTGDMGTGDPYEHRYAPSTPLAQLPFAVGLAHARKKDDVVALAVMGDGTSGEDFHEALNFPAVRAPVFLVQNNGRAISVPLDKOTAAPSLAKAVGVMGRILVDGDAVAHVHVSIDNARHARAGCG"

1..35824 /note="Nominal overlap with Streptomyces coelicolor cosmid GD3."

1..35824 /complement(2..622)

1..35824 /gene="bkdA2"

1..35824 /note="Pfam match to entry PF00676 El.dehydrog. Dehydrogenase El component, score 109.40, E-value 7.1e-29."

1..35824 /gene="SCH69.02"

1..35824 /gene="SCH69.02"

1..35824 /note="SCH69.02, possible transcriptional regulator. len: 135aa; similar to many eg. SW:LFP-ECOLJ leucine-responsive regulatory protein from Escherichia coli (163 aa) (fasta scores: opt: 181, z-score: 249.4, E(): 1.5e-06, (29.5% identity in 129 aa overlap). Contains Pfam match to entry. PF01037 ASNC-trans-reg. Asnc family and a probable helix-turn-helix 9-30aa (+3.69 SD)."

1..35824 /codon_start=1

1..35824 /transl_table=11

1..35824 /product="putative transcriptional regulator"

1..35824 /db_xref="GI:5102784"

1..35824 /translation="MQADGRASISVAERHYVSAANAYARINRLVEGVYRGFARYDERRGHGTSAYITIKIVNSWRTVRAQLRQLPQASHALAVGGDFVLLVHTPPDNNRA


```

/codon_start=1
/product="plasma membrane intrinsic protein"
/protein_id="AAD28761.1"
/db_xref="gi:4741931"
/translation="MARDIEAALAAHEGKDYSDPPAPLVDAEELTKSLYRAVIAEEV
ATLEFLIVATVIGYKHODPAAGPDAAGPAGVGLIAAFGGMIFLYCTAGVS
GGHINPAVTFGLFARVKSILVALLTYAOGLVAGVKGFGSAFYVYGGGANE
LSAGSKGTGLAETIGTFEVLVYVFSTDPKRNARSDSHVLAPELPGFAPVYHLA
TIPTGTGINPARSLGAALVYNNKAMDHWTFWGFPGALAAAHQVYLRASAK
LGSSASFSR"

```

```

BASE COUNT      228 a      381 c      357 g      241 t
ORIGIN
Query Match      3.3% Score 48.2; DB 8; Length 1207;
Best Local Similarity 47.2% Pred. No. 7.5;
Matches 146; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

```

```

OY 1023 GGACTACTACGACATCCATGACCTGCTCATATCCCAACCCAGTTTGACGGACGTTCCGA 1082
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 GGCCCTACACCTTCGGCCCTTCCTGCGCGCAAGGTCTCGCTGGCGCGGCTCTCTCTA 441
OY 1083 CTACGAGCCCTTTATCTGAACGCGATTCCGTCGCTGAGTCTTACGCGCGCGGAGG 1142
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 CATCGTGGCGAGTCCCTCGCGGCGCATCTCGCGGCTCGTCAAGGCTTCCAGAG 501
OY 1143 CATCATGCCAAGAGAACGACGCGCTGGGAGGTCAAGCCGCGCTGCGCTACGAGCG 1202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 502 CGCCTTCTACGCTGCGCTACGCGGCGGCGCAACAGTCAAGCGCGGCTACTCCAAAGG 561
OY 1203 CAATACCAACGCGCGGAGCAACATGACCAACCTCAACATGAGCCTTCCTGATCAA 1262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 562 CACGGCGCTCGCGCGGAGCATGATGGACACTTCTGCTCTACACCTCTTCTTCTCCG 621
OY 1263 CTCCAAGCCACCGCTTCGCGCTGCGCACTACGCCAAGCACTCTCTGATGCCAA 1322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 622 CACGACCCCAAGGCGACGCGCGTGTACTGATGTCCTGCTGCTGCTGCTCCAT 681
OY 1323 ACGGAATAC 1331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 682 CGGCTTTC 690

```

```

RESULT 13
SC9B5      36368 bp   DNA      BCT      13-JAN-1999
LOCUS      Streptomyces coelicolor cosmid 9B5.
DEFINITION
ACCESSION  AL035206.1 GI:4158179
VERSION     A035206.1 GI:4158179
KEYWORDS   ADA-like regulatory protein; aldehyde dehydrogenase; Arac;
            beta-galactosidase; cation-transporting ATPase; cold shock; dcdA;
            diaminomelate decarboxylase; gamma-glutamyl transferase; inositol
            monophosphatase; methylated-DNA-protein-cysteine methyltransferase;
            ogt2; oxidoreductase; Sir2 family; substrate binding protein;
            transport permease protein; tRNA synthetase.
SOURCE      Streptomyces coelicolor A3(2)
ORGANISM    Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomyces.
REFERENCE   1 (bases 1 to 36368)
AUTHORS     Seeger,K.J. and Harris,D.
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 36368)
AUTHORS     Bentley,S.D., Parkhill,J., Barrell,B.G. and Rajandream,M.A.
TITLE       Direct Submision
            Submitted (07-JAN-1999) Streptomyces coelicolor sequencing project,
            Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
            CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
            David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
            Colney, Norwich, Norfolk NR4 7UH, UK
            3 (bases 1 to 36368)
REFERENCE   Redenbach,M., Kleser,H.M., Denapalte,D., Eichner,A., Cullum,J.,
            Kinaehl,H. and Hopwood,D.A.

```

TITLE
JOURNAL
MEDLINE
COMMENT

A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome
Mol. Microbiol. 21 (1), 77-96 (1996)
97000351

NOTES:

Streptomyces coelicolor sequencing at The Sanger Centre is funded by the BBSRC.
Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web.
(URL: <http://www.sanger.ac.uk/Projects/S.coelicolor/>) CDS are numbered using the following system eg SC7B7.01c, SC (S. coelicolor), 7b7 (cosmid name), .01 (first CDS), c (complementary strand).

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS.

Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the Frameplot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.go.jp/jun/cgi-bin/frameplot.pl>. CAPTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, ttg or (atc)) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid 9B5 lies between IC3 and 9B1 on the AseI-B genomic restriction fragment.

FEATURES

source

1. 36368

/organism="Streptomyces coelicolor A3(2)"

/strain="A3(2)"

/db_xref="taxon:100226"

/clone="cosmid 9B5"

1. 1062

/gene="SC9B5.01"

<1. 1062

/note="SC9B5.01"

len: 353aa; overlaps with and forms the C-terminus of TR:069828 on overlapping cosmid 1A6, similar to several

eg. SW:PVD4_BURKE L-ornithine 5-monoxygenase from Burkholderia cepacia (444 aa) fasta scores: opt: 641,

z-score: 580.6, E(): 4.7e-25,(40.5% identity in 316 aa overlap).

/codon_start=1

/transl_table=11

/product="putative oxidoreductase"

/protein_id="CAA22743.1"

/db_xref="gi:4158180"

1. 102

/gene="SC9B5.01"

/note="Overlap with Streptomyces coelicolor cosmid 1A6 (EMBL:AL023496)."

1059. 1445

/gene="SC9B5.02"

1059. 1445

/gene="SC9B5.02"

1059. 1445

/gene="SC9B5.02"

1059. 1445

/gene="SC9B5.02"

1059. 1445

/gene="SC9B5.02"

1059. 1445

/gene="SC9B5.02"

1059. 1445

/gene="SC9B5.02"

1059. 1445

/gene="SC9B5.02"

1059. 1445

/gene="SC9B5.02"

440aa): similar tetramy eg. SM:DCDA_ECOLI_dcdA, diaminopimelate decarboxylase from *Escherichia coli* (420 aa) fasta scores: opt: 1769, z-score: 2033.9, E(): 0 (63.6% identity in 398 aa overlap). Contains p500878 Orn (DAP/Arg decarboxylases family 2 pyridoxal-P attachment site and p500879 Orn /DAP/Arg decarboxylases family 2 signature 2. Also contains Pflam match to entry PF00278 Orn_DAP_Arg-dec, Pyridoxal-dependent decarboxylase."

Orn_DAP_Arg-dec, Pyridoxal-dependent decarboxylase."

/codon_start=1
/trans_table=11
/product="diaminopimelate decarboxylase"
/protein_id="CAA22747.1"
/db_xref="GI:4158184"
/translation="MTSPSAPSSAERPRLPGALDGPRLAAAEHGPILMWYDAITIR
AIDRLRRPVIYRAKAGCSNLHILNREGGVADVASEEIRLALAGYRGVGGDD
PIVATRLRSTLRVYELGIPNALSPPKLDGVGAAPHPWIRINPFGHSHS
KNTNGEHSKHGITHHEHLEESLAVDHGDLVLELHMIGSVYGHLESYCEHYKK
VRNAGRIKRAISAGGGLSVPTPEDPELIDRTRVELMDAARELVSELGHPVRLIEE
GRFLVAGAVLAALAEVRAOKPVGSNYFVLVDAGFNDLKPAMYGSHRVSVDADGAPR
ASDRIDYVLVAGPLCESGDVFTOYGGGVPEVPVPRDTGDLVSHDGAAGASMSST
NSRLPIEVLVDGAETRLIRROTVAELLAPLEPFGALSPRSPRDR"
4415..5497
/gene="SC9B5_05"
/note="Pflam match to entry PF00278 Orn_DAP_Arg-dec,
pyridoxal-dependent decarboxylase."
4484..4540

misc_feature
misc_feature

Query Match 3.3%; Score 47.8; DB 1; Length 36368;
Best Local Similarity 52.2%; Pred. No. 6.6;
Matches 106; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 1116 CGGTGACTCTTCACGGGGCGCGAGGATCATGTCCGAAGAAGCAAGCCGTGGGG 1175
 ||||| ||||| ||||| ||| ||
Db 25978 CCGTGCACACTTCCCTCCACTTCGAGGGGTCAAGTCCCGCACTACGTCGTGATCAAGG 26037
 ||||| ||||| ||||| ||||| ||

QY 1176 AGGTCAAGCCGGCGGTGGCTACGACGCCAATCAACGCCGGGGGAGCAACATACCAA 1235
 ||||| ||||| ||||| ||||| ||

Db 26038 CGAACTGCGCGGCTACGACGAGAACTCTTACACGCCCTCCGATACGAGTACCGACCT 26097
 ||||| ||||| ||||| ||||| ||

QY 1236 CCTCAACCATGAAAGCCTTCTGTATCACTCCAAAGCCACCGCCTTGGCGTGCACACTA 1295
 ||||| ||||| ||||| ||||| ||

Db 26098 COTCAAGCCGGGGACCAACCAAGATGGCGGTGAGGTCTACCGCTACCTCGAGCGGAGCTG 26157
 ||||| ||||| ||||| ||||| ||

QY 1296 CGCCACGACCTCTCTCGATCC 1318
 ||||| ||||| ||||| ||||| ||

Db 26158 GATGGAGACCAAGACGACATGATCC 26180
 ||||| ||||| ||||| ||||| ||

RESULT 14
AC022472/c
LOCUS AC022472 92710 bp DNA PLN 15-FEB-2000
DEFINITION Arabidopsis thaliana chromosome 1 BAC T20H2 sequence, complete
sequence.
AC022472
AC022472.2 GI:6978398
HTG.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; euroids II; Brassicales; Brassicaceae;
Arabidopsids.
1 (bases 1 to 92710)
REFERENCE
AUTHORS Sakano,H., Yu,G., Lee,J.M., Lenz,C., Liu,S., Pham,P., Tortum,M.,
Chin,C., Chlouy,O., Choi,E., Chung,M., Gonzalez,A., Howmy,B.,
Liu,A., Vaysseu,M., Altairi,H., Brooks,S., Buehler,E., Chao,Q.,
Conn,L., Conway,A.B., Hansen,N.F., Johnson-Hopson,C., Khan,S.,
Kim,C., Lam,B., Miranda,M., Nguyen,M., Palm,C.J., Shima,P.,
Southwick,A., Davis,R.W., Ecker,J.R., Federspiel,N.A. and
Theologis,A.
Arabidopsis thaliana chromosome 1 BAC T20H2 sequence
Unpublished
2 (bases 1 to 92710)

AUTHORS	Theologis,A.
TITLE	Direct Submission
JOURNAL	Submitted (04-FEB-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA 3 (bases 1 to 92710)
REFERENCE	Theologis,A. Direct Submission
TITLE	Submitted (15-FEB-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA On Feb 15, 2000 this sequence version replaced gi:6899641.
COMMENT	
FEATURES	source location/Organism=Arabidopsis thaliana /cultivar=Columbia /db_xref=taxon:3702 /chromosome=1 /clone=T20H2
BASE COUNT	29482 a 17092 c 17067 g 29069 t
ORIGIN	
Query Match	3.3%; Score 47.6; DB 50; Length 92710;
Best Local Similarity	50.4%; Pred. No. 6.6;
Matches 116; Conservative	0; Mismatches 114; Indels 0; Gaps 0
Dn	1134 CGCGGAGGGATCATGTCCGAAGAAGCAACGCCCTTGCGGAGTCAAGCGCGGTGC 1193 Cc CCATCCTCCATCACACCACCTTAAGCCTTCACAAGCCAAAGCCAATGCGACACCGATGCC 33622
Oy	1194 CTACGAGCGCCAACTACCAAGCCCCGCGGAGAACATGACCAACTCAACATGAGACTT 1253 Ca CACCAACACTCTCTAATAACCTCAACGAGACGACGACACCGCCNAAACCAAAGCGAGCG 33562
Oy	1254 CCTGATCACTCCAAAGCCAGCCGCTTGGCGGTGCGCACCTTAGCCAAAGACCTTCCTC 1313 Cb CACCAACGAGCAACCAAGCCAGTCCCATGTCTCATACCAACTTAAGCCACCAAGCCCACCTC 33502
Dn	1314 GATCCCAAGCGANATACCATCTCTCTTGACAGCGAGCGCGGCACCA 1363 Db 33501 CAAAGCCTGTCCACACCATGTCTCAAAAACCTGCTCCACCA 33452
RESULT 15	
LOCUS	OPU75930/c OPU75930 bp DNA circular VRL 06-MAR-1998
DEFINITION	Orygia pseudotsugata nuclear polyhedrosis virus complete genome.
ACCESSION	U75930
VERSION	U75930.1 GI:2934903
KEYWORDS	Orygia pseudotsugata nuclear polyhedrosis virus. Orygia pseudotsugata nuclear polyhedrosis virus viruses; dsDNA viruses, no RNA stage; Baculoviridae; Nucleopolydnavirus.
SOURCE	1 (bases 1 to 13193) Ahrens,C.H., Russell,R.L., Funk,C.J., Evans,J.T., Harwood,S.H. and Rohrmann,G.F.
ORGANISM	Rohrmann,G.F. The sequence of the Orygia pseudotsugata multinucleocapsid nuclear polyhedrosis virus genome Virology 229 (2), 381-399 (1997)
TITLE	2 (bases 1 to 13193) Rohrmann,G.F. Direct Submission
JOURNAL	Submitted (23-Oct-1996) Oregon State University, Agricultural Chemistry, Corvallis, OR 97331-7301, USA
REFERENCE	3 (bases 1 to 13193) Rohrmann,G.F. Direct Submission
TITLE	Submitted (06-Mar-1998) Oregon State University, Agricultural Chemistry, Corvallis, OR 97331-7301, USA Sequence update by submitter On Mar 6, 1998 this sequence version replaced gi:1911246, similar to Autographa californica nuclear polyhedrosis virus
REMARK	
COMMENT	

FEATURES

source

(AcMNPV) complete genome: GenBank Accession Number L22898.

1. 133993 location/qualifiers

organism="Oryzia pseudotsugata nuclear polyhedrosis virus"

CDS

/db_xref="taxon:10450"

complement(123..947)

/note="ORF1, pk-1, AcMNPV ORF10 homolog"

/codon_start=1

/product="protein kinase 1"

/protein_id="AAC59000.1"

/db_xref="GI:1911247"

/translation="MDALQSLSEFVDAICAVIAKPVYNGRGKMDVLHHRPTSKFLF
RKTLAASFSADEINVDHSDHSPFADMEYCSPPATMLVYVCPDLFEIQQD
GALDNALVYINVRQLCDALNDLHNATITINDYKLENVLFQADARYLIDCYGLCKEE
HSPVQDGLTEFSEKIRHNINARSFEMWAVGLAYKLGLGRHPEFERSADVELDILAS
MRNNOYDNPDAALAKNNVIMARDFVCLTRFNEECSTSDYKQIAKHSFLASRHDI"
946..2370

/note="ORF2, AcMNPV ORF9 homolog"

/codon_start=1

/product="1628-capsid"

/protein_id="AAC59001.1"

/db_xref="GI:2934904"

/translation="MMEROYOSVRSYLINNOHNAIAAGPELQVAPCEAHSVGRNVD
RAVLRQAVLIDLKLAEDLYADATVYQADQPEASHPATILNMRLLIGVOPDAR
RINNSVARTAEILLRVYNDVADENYVSGPEFEYSKISYCOEFAQTPASASQNRQ
TSLPRPQTSIPROTOASLPQOTPPDQEMKSPSFYHTTRAILPQTOTPARITDFSRP
SDEVIYVGERAVPDIIRKPVPPKFEHLKSRSSVATNAAGATPAAPPPPPSAD
VTSMPPPPPPPSADVTSMPPPPVVDATSMPPPPPPPVVDATSMPPINNA
INNLIDAMVAETNNKAGDNRSALDIDKQGTLLKQPADGAPATPRSTLSEIQ
GTLTKLRKIEDQSSQTDLKDDVDTTKTILKNFTNIDIRISKQOEEMKRDJITIT
KRPPVHTDGNSTGNSDDMRD"

complement(2253..3370)

/note="ORF3, ph, AcMNPV ORF8 homolog"

/codon_start=1

/product="polyhedrin"

/protein_id="AAC59002.1"

/db_xref="GI:1911249"

/translation="MPDYSTRPTIGRTYVDNKKYKLGSYIKAKKKHLLHEEDD
KHLDPDHYVADDEPLFGPKNOKLTIEKIRANPKPTMKLIVWSKELREWTWV
VEDSFIVNDQVDEVDYFLVYNNRPTNRKYKLAQALAMDQDYVHEVIRIYEPY
VGNNNERYISLAKKGCCGCPIMINIAEYTNESFESVNVNVIWENFYKPIYVIGTDSSEEE
ELIEVSLVFKEVFEAPDAPLFTGPAP"

complement(3350..3541)

/note="ORF4"

/codon_start=1

/product="unknown"

/protein_id="AAC59003.1"

/db_xref="GI:1911250"

/translation="MRKNNGASQYPEHLTIVHKRHRQLSCWCMDDQSRNDPRAKGL
MEFHMFCGLRIYYOQIKH"

complement(3698..4093)

/note="ORF5"

/codon_start=1

/product="unknown"

/protein_id="AAC59004.1"

/db_xref="GI:1911251"

/translation="MGASQNSEQKILKLFYRMSQGAALDEKDLHCLYDERFVGA
HLKKRADDKARKKCAEKALKRLEIADHRALEAAAPACADGGRWSTLSRAOLDID
ARKEIYDRIRHQLKODSLKODRLKRR"

complement(4125..4739)

/note="ORF6, left-2, AcMNPV ORF6 homolog"

/codon_start=1

/product="late expression factor 2"

/protein_id="AAC59005.1"

/db_xref="GI:1911252"

/translation="MERVYVPAAGIDGLKRSYIYNDPHFVGLTSLPTVERGLE
VRNSGRMLALALAPKQEPQAPVRRPQPSRRNNVCLAKADAGQASLAKILARVSNP
PCSKSTIVADSSAPRGNNYKRRFEYCNQILANVITCYCKACGLIGLRBDACKV
GEVTHLLIKQDQYKPCSNCAKMKRYTKLCPQASGCKGLNICE"

complement(4742..4975)

/note="ORF7, AcMNPV ORF5 homolog"

/codon_start=1

CDS

...

CDS
/product="unknown"
/protein_id="AAC59006.1"
/db_xref="GI:1911253"
/translation="MNRPTMTAAVTTDYDREQLRREINSLRSVHELCTRSATGFD
CRFLFAGDRAPAVIYKAAANGCQSSSLCDKY"
5020..5460
/note="ORF8: AcMNPV ORF4 homolog"
/codon_start=1
/product="unknown"
/protein_id="AAC59007.1"
/db_xref="GI:1911254"
/translation="MPGARFVRSRLTDFEKENYVAHYDHLMLGLRLLDGKVSADY
RRGFLSRVIVASACMAVVOYVARDATIDMRHOTPIRVCNQCNAHADVAPRBDHS
IARILLAEAGAVLVIDHPDVTGETEEGVELLEVQRINAGDL"
complement(5669..6151)
/note="ORF9: ptp-2"
/codon_start=1
/product="protein tyrosine phosphatase 2"
/protein_id="AAC59008.1"
/db_xref="GI:1911255"
/translation="MTEASQIDEHVFGVYGDNEAMLFEINHAIASVILSDVA
PIRHLGLPVGDIHIVCEAPTCALPAMPALDYVVRIGEGKRVLYHGYAASR
SALAAYVLMRSROMPYODALNAVQSKRRVAINDHFRELATRCYSRFVNDLKPQIV
"
complement(6129..6791)
/note="ORF10: ptp-1, AcMNPV ORF1 homolog"
/codon_start=1
/product="protein tyrosine phosphatase 1"
/protein_id="AAC59009.1"
/db_xref="GI:1911256"
/translation="MLTSRMAVITYTRTSERNCGAMCTNGVRRROYHLPSPVARAM
PDRNHEYTACGAVIEGTRLLCFKVLNLEIEYTSDEDRTAASVLAHRLAVI
DILNTRITDGAQVMKGLIKIRVPGAVDDIVAEFIETVEFERCTMLVAV
HMTGLNRSGLVLCRYMWERLGVSPDALARETARKEIKERTNYLQDLARKHVGQ
PN"
6819..7814
/note="ORF11: AcMNPV ORF11 homolog"
/codon_start=1
/product="unknown"
/protein_id="AAC59010.1"
/db_xref="GI:1911257"
/translation="MSLSKILVYAVYGSYNLPHDRGESYLRIIVHEHLNTVYSN
ASCVRDIATARCLNSGLCEPVAROLDVSVAARLSAMFGDQATGICADMRALA
DIDRRAPLARVGRANTFALDAIDISDYNNOCIGITPRMHRPGSCGTAVADV
DPDIPADGWWHTKFCVLTIMLVACGAVPAGCATLRIYVAKHIGPNDEGCAPLIAA
VYGRCAIGREHFARHKTACMHILFQPMRNDLTPADERHPCGVKDFGRCKDYTD
LSTHADALYIHGTDROKNALFDLLCQVNASIDIDADCIQCVNKEYATQNKRYKM"
complement(8180..9142)
/note="ORF12: AcMNPV ORF13 homolog"
/codon_start=1
/product="unknown"
/protein_id="AAC59011.1"
/db_xref="GI:1911258"
/translation="MESWMEGWNNADEQVNAEFDEQAVRYRAYDQRAHSDLVRRDVE
RCHPVFKRYVLETAGRCQSVDFCKGKITSIDLQRNFDRQHRQINELVIGAP
PAQDSIGSLFAKXGIVOLLOQFSFANNEVLAAGAKHDDRNLDKTAIVANHV
KTLNNSDKFTSAHKSFKLEVAREQEPQRIQTLDTKINALQCAAPRTTAGVFEPR
DVKHPLHAEVGRVEDRGVQIARFGQEEHFRKRIEFEGGMVADVGRAPNPLA
VACIEEFANGGHKIRLRPKVIEVDCAVNAVAKDIVKAILKMT"
complement(9094..9825)
/note="ORF13: lcf-1, AcMNPV ORF14 homolog"
/codon_start=1
/product="late expression factor 1"
/protein_id="AAC59012.1"
/db_xref="GI:1911259"
/translation="MAPCKYTERVQMMMDALAYNDSRLAFMTDRPRVVAHNEFDS
AAOLFAYIVKNSIDVHYKPLEGGREVIYADFKQCADKELMLKVNGATAPMLEF
EGKEDAVORIMSGNGRFLIMLKFCGKFKMDAPKSLREHMEVFRQPAKLYSGDIBPG
STADCYRAVHNTIGDAREDLVRIYPPVDVDRVFCANAKQIRAPFSYNTKGGDYSRL
TOOLOIRINACGAGCAGGTPPTSK"
9893..11362
/note="ORF14: egt, AcMNPV ORF15 homolog"
/codon_start=1

CDS
/product="ecdysteroid UDP-glucosyltransferase"
/protein_id="AAC59013.1"
/db_xref="GI:1911260"
/translation="MVFLLILTLATGARAASILAVLPTRYASHVYRAYVALYK
NCHNVTYIKPOLLDYAVDECGRVADQKQKLVASVSSVFRKRGVADETVY
TADNMGILMEKDFDANVRYRSTRTDVAVVEAFADYALVFGHLEFPAPVIOI
APGYGLAIEFRRRAVRAHPLHYPTFGAALTFRGGLSEMRLLNEPELLARSDILL
KOGRKSTPTTROLNDYVOLLNLNHPYDNNRRYPVSVOYLGGIHLAQLPORDLA
PLERLINESVGVAVSFSGIDTNSIAEFLQMLDFTFANLNTYLMKYDDVAAS
VALPRNVAQWFSQTAVLRHRKNVAFETQGLQSSDALQARVPMCLPVMGQEFH
ARKLQEGFVARALDITAAVASAPOLDAIEVLADEAVARADIKRAVVEHDAADDEKA
VKFERVVKFNNDVWMPARSJKTAAANAYSDYFRFPL"
11547..12140
/note="ORF15: AcMNPV ORF16 homolog"
/codon_start=1
/product="unknown"
/protein_id="AAC59014.1"
/db_xref="GI:1911261"
/translation="METHQPISTYAPRGAVCAVRYVYTTTYSDSGNNEDRLTQ
IVAOQLRRLNFSKISQLORRVRNMGLIKKNSVIANLALITOKTKTHTPATIR
KNVHTTSSSDQFQRQVLELCANGGEQVFCARADCAARRRRVAELATLALGAVVA
SAANRFEIIEDEKLYSAKLIVQOYLHDGDSHDCAD"
12109..12732
/note="ORF16: AcMNPV ORF17 homolog"
/codon_start=1

CDS
complement(6129..6791)
/note="ORF10: ptp-1, AcMNPV ORF1 homolog"
/codon_start=1
/product="protein tyrosine phosphatase 1"
/protein_id="AAC59009.1"
/db_xref="GI:1911256"
/translation="MLTSRMAVITYTRTSERNCGAMCTNGVRRROYHLPSPVARAM
PDRNHEYTACGAVIEGTRLLCFKVLNLEIEYTSDEDRTAASVLAHRLAVI
DILNTRITDGAQVMKGLIKIRVPGAVDDIVAEFIETVEFERCTMLVAV
HMTGLNRSGLVLCRYMWERLGVSPDALARETARKEIKERTNYLQDLARKHVGQ
PN"
6819..7814
/note="ORF11: AcMNPV ORF11 homolog"
/codon_start=1
/product="unknown"
/protein_id="AAC59010.1"
/db_xref="GI:1911257"
/translation="MSLSKILVYAVYGSYNLPHDRGESYLRIIVHEHLNTVYSN
ASCVRDIATARCLNSGLCEPVAROLDVSVAARLSAMFGDQATGICADMRALA
DIDRRAPLARVGRANTFALDAIDISDYNNOCIGITPRMHRPGSCGTAVADV
DPDIPADGWWHTKFCVLTIMLVACGAVPAGCATLRIYVAKHIGPNDEGCAPLIAA
VYGRCAIGREHFARHKTACMHILFQPMRNDLTPADERHPCGVKDFGRCKDYTD
LSTHADALYIHGTDROKNALFDLLCQVNASIDIDADCIQCVNKEYATQNKRYKM"
complement(8180..9142)
/note="ORF12: AcMNPV ORF13 homolog"
/codon_start=1
/product="unknown"
/protein_id="AAC59011.1"
/db_xref="GI:1911258"
/translation="MESWMEGWNNADEQVNAEFDEQAVRYRAYDQRAHSDLVRRDVE
RCHPVFKRYVLETAGRCQSVDFCKGKITSIDLQRNFDRQHRQINELVIGAP
PAQDSIGSLFAKXGIVOLLOQFSFANNEVLAAGAKHDDRNLDKTAIVANHV
KTLNNSDKFTSAHKSFKLEVAREQEPQRIQTLDTKINALQCAAPRTTAGVFEPR
DVKHPLHAEVGRVEDRGVQIARFGQEEHFRKRIEFEGGMVADVGRAPNPLA
VACIEEFANGGHKIRLRPKVIEVDCAVNAVAKDIVKAILKMT"
complement(9094..9825)
/note="ORF13: lcf-1, AcMNPV ORF14 homolog"
/codon_start=1
/product="late expression factor 1"
/protein_id="AAC59012.1"
/db_xref="GI:1911259"
/translation="MAPCKYTERVQMMMDALAYNDSRLAFMTDRPRVVAHNEFDS
AAOLFAYIVKNSIDVHYKPLEGGREVIYADFKQCADKELMLKVNGATAPMLEF
EGKEDAVORIMSGNGRFLIMLKFCGKFKMDAPKSLREHMEVFRQPAKLYSGDIBPG
STADCYRAVHNTIGDAREDLVRIYPPVDVDRVFCANAKQIRAPFSYNTKGGDYSRL
TOOLOIRINACGAGCAGGTPPTSK"
9893..11362
/note="ORF14: egt, AcMNPV ORF15 homolog"
/codon_start=1

Query Match 3.3%; Score 47.6; DB 16; Length 131993;
Best Local Similarity 46.0%; Pred. No. 6.4;
Matches 161; Conservative 0; Mismatches 189; Indels 0; Gaps 0;

QY 989 CGGCGGTCGCGCCAGATCGAGAAACTGTGAGAGCTACTACGACCTCCAGCTCC 1048
DB 75728 CGAGCGCGAGCGCAACGCCGACGCCAGCGCCAGCGCATCGCCGAGC 75669
QY 1049 CTCATATCCCAACCCAGTTTGACGGACGTTCCGACTACGAGGCTTTATCTGAAGGCA 1108
DB 75668 CCGAGCGCAGCGCTATACCAACGCCGAGCGCATCGCCGACCTATCAACAGCGCATCA 75609
QY 1109 TTCGTCGCGGTGACTCTTCACGGGGCGCGAGGCGCATCATGTCGAAGAAAGCAAGCC 1168
DB 75608 CCAACGCCGACCCATCGCGCGAGCGCATCGCGAGCGCATCAACAGCGCGCATCG 75549
QY 1169 GGTGGGAGGTCAAGCGCGCGGCGCTACGACCACTACCAAGCGCGGGGAGACACA 1228
DB 75548 CCGAGCGCATCAACAGCGCGCGCATCGCGAGCGCATCAACAGCGCGCGCATCA 75489
QY 1229 TGACCAACCTCAACCATGAGGCTTCCTGATCACTCCAAAGCAGCGGCTTCGCGCTCG 1288
DB 75488 CCAAGCGCGAGCGCATCGCGCGAGCGCATCAACAGCGCATCAACAGCGCGCATCG 75429
QY 1289 CCACTACGCCCAAGCACTCTCTCGATCCCAAGGAAATCCCATCC 1338
DB 75428 CCGAGCGCATCAACAGCGCGCATCGCGAGCGCATCAACAGCGCGCATCG 75379

Search completed: May 10, 2000, 22:26:23
Job time: 4257 sec

D	b	166	GGTGGTAAAGCCCAAGCAAGACAGGTTAACTATCTCTACGAGGAGCTGAAGAAAGCTGGC	225
O	y	161	TACTATGATGTCTCAACAGACACCTCAGGTGCACCTGTGGAGCAATGCCAGACACGCTC	240
D	b	226	TACTATGATGTCTCAACAGACACCTCAGGTGCACCTGTGGAGCAATGCCAGACACGCTC	285
O	y	241	AAGGTGGGCGATGAGAAATCGAAGGCAAAACCATGACCTACAGTCCGACGCTCGAGGTC	300
D	b	286	AAGGTGGGCGATGAGAAATCGAAGGCAAAACCATGACCTACAGTCCGACGCTCGAGGTC	345
O	y	301	ACCGCCGATGTAGCCGCTGCTCAAGAACCTTGGATGCAAGCGAGGCGGATTAACCATCCGAT	360
D	b	346	ACCGCCGATGTAGCCGCTGCTCAAGAACCTTGGATGCAAGCGAGGCGGATTAACCATCCGAT	405
O	y	361	GTGAGGGGCAAGGTGCGCCGTGATCAACCGTGGAGTAATGCCCGTGTGGGCAAGTCCGTT	420
D	b	406	GTGAGGGGCAAGGTGCGCCGTGATCAACCGTGGAGTAATGCCCGTGTGGGCAAGTCCGTT	465
O	y	421	CTCGCTGCCAAAGCCAAAGCGCCGCGCTTCGATTTGTCTATAACATGTGTGGCCGGATCCATG	480
D	b	466	CTCGCTGCCAAAGCCAAAGCGCCGCGCTTCGATTTGTCTATAACATGTGTGGCCGGATCCATG	525
O	y	481	GCGGGCAACCTTGGCGGGGGGCGAGAGTGAATAAGGACCGTATTCGGCCATTTCGGTATC	540
D	b	526	GCGGGCAACCTTGGCGGGGGGCGAGAGTGAATAAGGACCGTATTCGGCCATTTCGGTATC	585
O	y	541	AGCTTGAGAGATGTGGCCAGAGCTGATCAAGCTTGTCTGAGAGCTGATTCGGTATCTGGAT	600
D	b	586	AGCTTGAGAGATGTGGCCAGAGCTGATCAAGCTTGTCTGAGAGCTGATTCGGTATCTGGAT	645
O	y	601	CTGTGGGTGGATTAAGCAGAGAGAACCGTAGACGTATTAACGTTGTGGCCAGACGAG	660
D	b	646	CTGTGGGTGGATTAAGCAGAGAGAACCGTAGACGTATTAACGTTGTGGCCAGACGAG	705
O	y	661	GGCGGCGATCCGAACAACGTCGTGGCCGTGGGGGCGCACACGAGACATCGAGCGGGG	720
D	b	706	GGCGGCGATCCGAACAACGTCGTGGCCGTGGGGGCGCACACGAGACATCGAGCGGGG	765
O	y	721	CTGTGATCAACGACGATGGCTGGGGCATTTATAGCACTTGATGCTATTCGCAAGCGCTC	780
D	b	766	CTGTGATCAACGACGATGGCTGGGGCATTTATAGCACTTGATGCTATTCGCAAGCGCTC	825
O	y	781	ACGGATCTCCGACGAAGATGGCGCGCTCTCTTGTGGACAGCAGAGAGTTCCGT	840
D	b	826	ACGGATCTCCGACGAAGATGGCGCGCTCTCTTGTGGACAGCAGAGAGTTCCGT	885
O	y	841	CTGCTGGGCAACACTACTACGTCTCCCATCTGAATGCCACCGAGCTGAACAAGATCCGA	900
D	b	886	CTGCTGGGCAACACTACTACGTCTCCCATCTGAATGCCACCGAGCTGAACAAGATCCGA	945
O	y	901	CTGTACCTGAACCTTCGACATGATGGCTCACTAATTAAGCCCTCATATCTATAGACGT	960
D	b	946	CTGTACCTGAACCTTCGACATGATGGCTCACTAATTAAGCCCTCATATCTATAGACGT	1005
O	y	961	GATGGATCGGCTTCAACCGAGCGGACCGCGGCTTCCGCCAGATCGAGAAATGTTTC	1020
D	b	1006	GATGGATCGGCTTCAACCGAGCGGACCGCGGCTTCCGCCAGATCGAGAAATGTTTC	1065
O	y	1021	GAGGACTACTACGACTCCATCGACCTGCTCATATCCCAACCGAGTTTGAAGGAGTTTC	1080
D	b	1066	GAGGACTACTACGACTCCATCGACCTGCTCATATCCCAACCGAGTTTGAAGGAGTTTC	1125
O	y	1081	GACTACGAGGCTTTATCTCGAAACGGATTCGCTCGCGGTGGAGCTCTTCAACGGGGCGGAG	1140
D	b	1126	GACTACGAGGCTTTATCTCGAAACGGATTCGCTCGCGGTGGAGCTCTTCAACGGGGCGGAG	1185
O	y	1141	GGCATCATGTCCGAAGGAACGCAAGCCGTGGGAGGTCAAGCCGCGGTGGCTTACAC	1200
D	b	1186	GGCATCATGTCCGAAGGAACGCAAGCCGTGGGAGGTCAAGCCGCGGTGGCTTACAC	1245
O	y	1201	GGCAATCAACAGCGCGGGGGAAGAACATGACCAACTTAACCATGAAGCTTCTGATC	1260

DB	1246	GCCACCTACACAGCGCGGAGACAAACATGACCAACCTCAACCATGAAAGCCTTCTGATC	1305
QY	1261	AACTCGAAGACACACCGGCTTGGCCCTGCGACACTACGCGCAAGACTCTCTCGATCCCC	1320
DB	1306	AACTCGAAGACACACCGGCTTGGCCCTGCGACACTACGCGCAAGACTCTCTCGATCCCC	1365
QY	1321	AAACGGAATACACATCTCTCTGACACGACGAGCCCGACCATCGACATTCGGCAAG	1380
DB	1366	AAACGGAATACACATCTCTCTGACACGACGAGCCCGACCATCGACATTCGGCAAG	1425
QY	1381	AGAGTCCGAGACACACAGCTCAGTATCAGGATCCGGATGCTGGCATTTCTAAGTCGAG	1440
DB	1426	AGAGTCCGAGACACACAGCTCAGTATCAGGATCCGGATGCTGGCATTTCTAAGTCGAG	1485
QY	1441	GCA 1443	
DB	1486	GCA 1488	
RESULT	2		
ID	V82521		
AC	V82521 standard; DNA; 1491 BP.		
DE	17-MAR-1999 (first entry)		
DE	Aspergillus oryzae aminopeptidase II encoding DNA.		
KW	Aminopeptidase; protein hydrolystase; glycine releasing; protease;		
KW	proteinaceous material; flavour; food; baking; animal feed additive;		
KW	palatability; hydrolysis; solubility; emulsifying; foaming; aroma;		
OS	mouthfeel; crust colour; ss.		
KS	Aspergillus oryzae.		
FM	Key	Location/Qualifiers	
FT	CDS	1..1491	
FT		/*tag= a	
PN	W09851163-A2.		
PD	19-NOV-1998.		
PF	15-MAY-1998; U09998.		
PR	16-DEC-1997; US-069719.		
PR	16-MAY-1997; US-857886.		
PR	20-OCT-1997; US-062893.		
PR	16-DEC-1997; DK-001465.		
PA	(ASAH) ASAH CHEM IND CO LTD.		
PA	(NOVO) NOVO NORDISK BIOTECH INC.		
PI	Blinskovsky A, Brown K, Byun T, Fujii M, Gollightly E,		
PI	WPI: 99-045177/04.		
DR	P-PSDB: W89386.		
PT	Production of protein hydrolysate - using protease and enzyme that		
PT	releases glycine, useful as flavour improvers in foods and animal		
PT	feed additives		
PS	Claim 4: Fig 1: 84bp: English.		
CC	A method has been developed for the production of protein hydrolysates		
CC	(PH) comprising reacting a protein with: (i) at least one polypeptide		
CC	with gly-releasing activity; and (ii) at least one other protease so		
CC	that the amount of Gly produced is greater than when the protease is		
CC	used alone. PH are used to improve flavour of foods (e.g. baked goods)		
CC	and as animal feed additives. PH are preferably also enriched in glu		
CC	(free and/or peptide bound), so have improved flavour and palatability.		
CC	Addition of a polypeptide with Gly-releasing activity increases the		
CC	degree of hydrolysis (or reduces the amount of enzyme needed) and		
CC	hydrolysates have better solubility, and emulsifying and foaming		
CC	properties. Baked goods containing them have improved aroma, mouthfeel		
CC	and crust colour. The present sequence encodes Aspergillus oryzae		
CC	aminopeptidase II, which is used in the method of the invention.		
CC	Sequence 1491 BP; 335 A; 438 C; 426 G; 292 T;		
QY	Query Match	100.0%; Score 1443; DB 1; Length 1491;	
	Best Local Similarity	100.0%; Pred. No. 0;	
	Matches 1443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
	1 GGGAGGCGGCTTTGTCGCCGAGTACATCCCCGAGGATATTCAGTTGGAAGATCTGCTG 60		

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GGGAGGCGCTGTTTCGCCGAGATAGTCCCGAGATATTCACTGGAGATCTGCTG 60
Db 46 GGGAGGCGCTGTTTCGCCGAGATAGTCCCGAGATATTCACTGGAGATCTGCTG 105
QY 61 GAAGATATCCCAACACTGAGGACTTCGCTATGCTACCCGAGCGCAATGGCTCTT 120
Db 106 GAAGATATCCCAACACTGAGGACTTCGCTATGCTACCCGAGCGCAATGGCTCTT 165
QY 121 GGTGTAAAGCCACGAGACAGCGTTAACTATCTCTACGAGAGCTGAAGAAGTGGC 180
Db 166 GGTGTAAAGCCACGAGACAGCGTTAACTATCTCTACGAGAGCTGAAGAAGTGGC 225
QY 181 TACTATGATGTCTACAGACAGCTCAGGTGACCTGTGGAGCAATGCCAGACGCTC 240
Db 226 TACTATGATGTCTACAGACAGCTCAGGTGACCTGTGGAGCAATGCCAGACGCTC 285
QY 241 AAGTGGGCGANTGAGAAATCGAGGCGAAGACATGACTACATGCTCCAGCGTGGAGTC 300
Db 286 AAGTGGGCGANTGAGAAATCGAGGCGAAGACATGACTACATGCTCCAGCGTGGAGTC 345
QY 301 ACCGCGATGTAGCCGTGCTCAAGAACCTGGGATGACGAGGCGGATTACCATCGAT 360
Db 346 ACCGCGATGTAGCCGTGCTCAAGAACCTGGGATGACGAGGCGGATTACCATCGAT 405
QY 361 GTCGAGGCGAAGGTGCGCTGATCAAGGTGGAGATGCCCTTGGCGACAATCGCTT 420
Db 406 GTCGAGGCGAAGGTGCGCTGATCAAGGTGGAGATGCCCTTGGCGACAATCGCTT 465
QY 421 CTCGCTGCAAAAGCCAGGCGCGCTTCGATTGTCTATTAACATGTGGCGGATCATG 480
Db 466 CTCGCTGCAAAAGCCAGGCGCGCTTCGATTGTCTATTAACATGTGGCGGATCATG 525
QY 481 GCGGCGACCTTGGCGCGCGCGAGATAGAGGACCGATTTCGCCATTCGCGATC 540
Db 526 GCGGCGACCTTGGCGCGCGCGAGATAGAGGACCGATTTCGCCATTCGCGATC 585
QY 541 AGCTTGGAGATGGCCAAAGCTGATCAAGCTTGTGAGGCTGGATTCGATCTGTGAT 600
Db 586 AGCTTGGAGATGGCCAAAGCTGATCAAGCTTGTGAGGCTGGATTCGATCTGTGAT 645
QY 601 CTGTGGTGGATAGTAACAGAGAGACGTAACGATTAACGTTGCGCGAGACGAG 660
Db 646 CTGTGGTGGATAGTAACAGAGAGACGTAACGATTAACGTTGCGCGAGACGAG 705
QY 661 GCGGCGATCCGAACACGTCGCTGCGTGGGTGGCCACACGACTCAATCGAGGCGGC 720
Db 706 GCGGCGATCCGAACACGTCGCTGCGTGGGTGGCCACACGACTCAATCGAGGCGGC 765
QY 721 CCTGTATCAACAGACGATGGCTCGGGCATTTATACCACTTGGTATGGCCAAACGCTC 780
Db 766 CCTGTATCAACAGACGATGGCTCGGGCATTTATACCACTTGGTATGGCCAAACGCTC 825
QY 781 ACGGAGTACTCCGTCAAGAAATGCGTGGCTTCCCTTCTGGACAGAGAGTTCGCT 840
Db 826 ACGGAGTACTCCGTCAAGAAATGCGTGGCTTCCCTTCTGGACAGAGAGTTCGCT 885
QY 841 CTGCTGGGACGACACTAGCTCTCCATCTGAATGCCACCGAGCTGAACAAATCCGA 900
Db 886 CTGCTGGGACGACACTAGCTCTCCATCTGAATGCCACCGAGCTGAACAAATCCGA 945
QY 901 CTGTACCTGACCTTCGACATATATGCTCCCTACCTAATACGCGCTATATCATAGCGT 960
Db 946 CTGTACCTGACCTTCGACATATATGCTCCCTACCTAATACGCGCTATATCATAGCGT 1005
QY 961 GATGATGAGGCTTCAACAGAGCGAGCGGCGGTTCCGCCACAGATGAGAAATGTTTC 1020
Db 1006 GATGATGAGGCTTCAACAGAGCGAGCGGCGGTTCCGCCACAGATGAGAAATGTTTC 1065
QY 1021 GAGGACTACTACGACTCGACCTGCTCATATATCCCAACCCAGTTTACGAGCGTTCC 1080
```

```
Db 1086 GAGGACTACTACGACTCGACCTGCTCATATATCCCAACCCAGTTTGAAGAGCTTCC 1125
QY 1081 GACTACGAGGCTTTATTCCTTAACGCGATTCCTCCGCTGAGCTTTCAGGCGCCGAG 1140
Db 1126 GACTACGAGGCTTTATTCCTTAACGCGATTCCTCCGCTGAGCTTTCAGGCGCCGAG 1185
QY 1141 GGCATCATGTCCGAAGAAAGCAAGCGGCTGGGAGAGTCAAGCGGCGTGGCTTACGAC 1200
Db 1186 GGCATCATGTCCGAAGAAAGCAAGCGGCTGGGAGAGTCAAGCGGCGTGGCTTACGAC 1245
QY 1201 GCCAATACACAGCGCGGAGAGACATGACCAACCTCAACATGAAAGCTTCTGATC 1260
Db 1246 GCCAATACACAGCGCGGAGAGACATGACCAACCTCAACATGAAAGCTTCTGATC 1305
QY 1261 AACTCCAAAGCCACCGCTTGGCGTGGCCACTACGCAACGACTTCTCGATCCC 1320
Db 1306 AACTCCAAAGCCACCGCTTGGCGTGGCCACTACGCAACGACTTCTCGATCCC 1365
QY 1321 AAAGGAATACACATCCTCCTTGCACGAGAGCGCGACCATGAGCATTTGGGCAAG 1380
Db 1366 AAAGGAATACACATCCTCCTTGCACGAGAGCGCGACCATGAGCATTTGGGCAAG 1425
QY 1381 AGAGCTCCGAAGACACAGCTCAGCTATCAGAGATCCGATGCTGGATTCGAAGTGAG 1440
Db 1426 AGAGCTCCGAAGACACAGCTCAGCTATCAGAGATCCGATGCTGGATTCGAAGTGAG 1485
QY 1441 GCA 1443
Db 1486 GCA 1488
```

RESULT 4

X34119
ID X34119 standard; DNA; 1503 BP.
AC X34119;
DT 06-JUL-1999 (first entry)
DE Mycobacterium species nucleic acid sequence 19D.
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW hybridisation; detection; vaccine; immunisation; infection; ss.
OS Mycobacterium sp.
PN W09J0186-A2.
PD 25-FEB-1999.
PF 14-AUG-1998; F01813.
PR 11-SEP-1997; FR-011325.
PR 14-AUG-1997; FR-010404.
PA (INSP.) INST PASTEREUR.
PI Guisqueno A; Lim EM, Pellicle V, Portnoi D, Gouquet de la Salmoniere Y,
PI Guisqueno A;
DR WPI: 99-181045/15.
DR P-PSDB: Y04867.
PT Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in
PT infection-associated protein expression
PS Claim 22; Fig 19D; 309pp; French.
CC Sequence X34001-X34252 represent nucleic acids encoding secreted
CC proteins from various Mycobacterium species microorganisms. The
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
SQ Sequence 1503 BP; 275 A; 470 C; 502 G; 256 T;

Query Match 12.3%; Score 178.2; DB 1; Length 1503;

Best Local Similarity 56.2%; Pred. NO. 3.5e-33;

Matches 386; Conservative 0; Mismatches 283; Indels 18; Gaps 2;

```
QY 640 AACCTTGTGCGAGACAGAGGCGCGATCCGAACAGCTGTCGCGTGGTGCCAC 699
Db 754 AACCTTGTGCGAGAGAGAGAGCGGCTGTCGCGCAAGTGTGATGAGAGTCCCAT 813
QY 700 ACGGACTCAGTGAAGCGGCGCTGTATCAACAGCAGATGGCTCGGCAATATTAGCAAC 759
```

Db 814 TTGGACAGGTTCCGGAGAGACCCGGATCAACAGACAGGCTCGGGAGTGTGCTGGGCT 873
Qy 760 TTGGTATTGCAAGAGCGCTACGCACTACTCC---GTCAAGATGCGCTGCGCTTCTC 816
Db 874 CTGGAAACCGCACTGACGATGGGGAACTCACCCATGATGTCCAAACGGGTACGGTGGCC 933
Qy 817 TTCTGACAGCAAGAGAGTTCGCTGCTGGGAGACAGCACTACAGCTCTCCATCTGAAT 876
Db 934 TTCTGGGGGCGCCGAGGATTCGGCTGATTTGGTTCACAGAACTACGTCGATCGCTGAGC 933
Qy 877 GCCACGAGCTGACAGAGATCCGACTGTACCTCAACTTCGACATGATCGCTCACTAAC 936
Db 994 ATGAGAGCGCTCAAGGCACTCGGCTGTATCTGACCTGACATGTTGGCTGGCCGAAAC 1053
Qy 937 TACGCCCTATGATCTATGACGCTGATGATCGCGC-----TTCAACAG 981
Db 1054 CCGGGTACTTCACTACAGCAGCGGTGACCAAGTCCGCTAGACGCCCGCGGTCAAGCCG 1113
Qy 982 AGGGAACCGCGGCTCCGCGCCAGATCGAGAGAACTGTGAGAGACTACTACGCTCATC 1041
Db 1114 GTGGTCCCGAAGGCTCGGCGGATGACGACGCTTCGCTGCTATCAGAAATGCGC 1173
Qy 1042 GACCTGCTCATATCCCAACCCAGTTGACGAGCTTCGACTACGAGGCTTTATCTG 1101
Db 1174 GGCAGAACCGCGAGAGACACTGCTTCGACGCTGCGTCCAGTACGAGGCTTCAGGCTG 1233
Qy 1102 AACGCAATTCGTCGGGTGAGCTCTTCACGGGCGCGAGGGCATCATGTTCGAGAGAAC 1161
Db 1234 GCGGGTATCCCTTCCTGCGGCTGTCTCCGGCGCTGAGTCAAGTATCGCGCGACAA 1293
Qy 1162 GCAAGCGCGGGGGAGGTCAAGCGCGGCTGAGAGCCAGCAACTACCGCGCGGGGA 1221
Db 1294 GCGGAGCTCGGGGCGGACCGCGAGAGACCTTTGATCCCACTTCACAGAGAAC 1353
Qy 1222 GACACATGACCAACTCAACCATGAGCCCTTCGATCACTCAACAGCCACCGCTTC 1281
Db 1354 GACACCTGAGCAATTCAGACCGACCGCGCTGGTATCAAGCGGCTGCGGTAC 1413
Qy 1282 GCGGTGCGCACTACGCCAAGCACTC 1308
Db 1414 GCGGTGGGTTGTATGCGCAGGACCTC 1440

RESULT 5
X34120
ID X34120 standard; DNA; 1587 BP.
AC X34120;
DE 06-JUL-1999 (first entry)
DE Mycobacterium species nucleic acid sequence 19F.
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW hybridisation; detection; vaccine; immunisation; infection; ss.
OS Mycobacterium sp.
PN WO909186-A2.
PD 25-FEB-1999.
PE 14-AUG-1998; F01813.
PR 11-SEP-1997; FR-011325.
PR 14-AUG-1997; FR-010404.
PA (INSP) INST PASTEUR.
PI Gicquel B, Lim EM, Pellicic V, Portnoi D, Goguet de la Salmoniere Y,
PI Guilguel A;
DR WPI: 99-181045/15.
DR P-PSDB; Y04868.
PT Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in
PT infection-associated protein expression
PS Claim 22; Fig 19F; 309pp; French.
CC Sequences X34001-X34252 represent nucleic acids encoding secreted
CC proteins from various Mycobacterium species microorganisms. The
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
SQ Sequence 1587 BP; 288 A; 490 C; 533 G; 276 T.

Query Match 12.3%; Score 178.2; DB 1; Length 1587;
Best Local Similarity 56.2%; Pred. No. 3.5e-33;
Matches 386; Conservative 0; Mismatches 283; Indels 18; Gaps 2;
Qy 640 AACGTGTGGCGAGAGAGGCGGCGATCCGAAACAGGTGCTGCGGTGGTGGCCAC 699
Db 838 AACGTATGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 897
Qy 700 ACAGACTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 759
Db 898 TTGGACAGGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 957
Qy 760 TTGGTATTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 816
Db 958 CTGGAAACCGCACTGACGATGGGGAACTCACCGCATGTCTCCAAACCGGTACGGTTCGC 1017
Qy 817 TTCTGACAGCAAGAGAGTTCGCTGCTGGGAGAGCACTACTAGCTCTCCATCTGAAT 876
Db 1018 TTCTGGGGGCGCGAGGAAATTCGCTGATTTGGTTCAGAACTACGATCGCTGAGC 1077
Qy 877 GCCACGAGCTGACAGAGATCCGACTGTACTGAACTTCGACATGATCGCTCACTAAC 936
Db 1078 ATGAGAGCGCTCAAGGCACTCGGCTGTATCTGAACTTGGACATGTTGGCTGGCGAAC 1137
Qy 937 TACGCCCTATGATCTATGACGCTGATGATCGCGC-----TTCAACAG 981
Db 1138 CCGGGTACTTCACTACAGCAGCGGTGACCAAGTCCGCTAGACGCCCGCGGTCAAGCCG 1197
Qy 982 AGGGAACCGCGGCTCCGCGCCAGATCGAGAGAACTGTGAGAGACTACTACGCTCATC 1041
Db 1198 GTGGTCCCGAAGGCTCGGCGGATGACGACGCTTCGCTGCTATGTAGATGCGC 1257
Qy 1042 GACCTGCTCATATCCCAACCCAGTTGACGAGCTTCGACTACGAGGCTTTATCTG 1101
Db 1258 GGCAGAACCGCGAGAGACACTGCTTCGACGCTGCGTCCAGTACGAGGCTTCAGGCTG 1317
Qy 1102 AACGCAATTCGTCGGGTGAGCTCTTCACGGGCGCGAGGGCATCATGTTCGAGAGAAC 1161
Db 1318 GCGGGTATCCCTTCCTGCGGCTGTCTCCGGCGCTGAGTCAAGTATCGCGCGACAA 1377
Qy 1162 GCAAGCGCGTGGGGAGGTCAAGCGCGGCTGAGAGCCAGCAACTACAGCGCGGGGA 1221
Db 1378 GCGGAGCTTCGGGCGGACCGCGAGAGCTTTGATCCCACTATACAGAGAAC 1437
Qy 1222 GACACATGACCAACTCAACCATGAGCCCTTCGATCACTCAACAGCCACCGCTTC 1281
Db 1438 GACACCTGAGCAATTCAGACCGGACCGGCTGGTATCAAGCGGCTGCGGTAC 1497
Qy 1282 GCGGTGCGCACTACGCCAAGCACTC 1308
Db 1498 GCGGTGGGTTGTATGCGCAGGACCTC 1524

RESULT 6
Q21833
ID Q21833 standard; DNA; 390 BP.
AC Q21833;
DE 08-JUN-1992 (first entry)
DE Randomising oligonucleotide used in SPERT mRNA prep.
KW Systematic polypeptide evolution by reverse translation; SPERT;
KW ligand binding; ss.
OS Synthetic.
PN WO9202536-A.
PD 20-FEB-1992.
PE 01-AUG-1991; U05463.
PR 02-AUG-1990; US-561968.
PA (CONS) UNIV OF COLORADO.
PI Gold L, Tuerk C;
DR WPI: 92-080018/10.
PT New method of systematic polypeptide evolution by reverse
PT translation - by linking each polypeptide in sample mixt. to

CC The identification of the sequences encoding the enzymes involved
CC in the biosynthetic pathway means that they can be isolated and
CC manipulated. Mutant microorganisms in which a step in the
CC streptomycin biosynthetic pathway is blocked can be cultured to
CC produce streptogramin intermediates, which may later be converted
CC to streptogramin derivatives. Recombinant cells may also be used
CC for the bioconversion of streptogramins from one form to another or
CC for the production of hybrid antibiotics.
SQ Sequence 833 BP; 106 A; 395 C; 250 G; 82 T;

Query Match 3.3%; Score 47.8; DB 1; Length 833;
Best Local Similarity 46.5%; Pred. No. 0.013;
Matches 154; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 984 GCGACGGCCGGTTCGGCCAGATGAGAAAGTGTTCGAGGACTTACGATCCATGGA 1043
DB 174 GCGCGTGGCCGCGGACGACGACATCGGCTGTATCAACGCCGCCGCTCCGCGGACCA 233
QY 1044 CTGCTCTATATCCCAACCACTTGTAGGAGAGCTTCCGACTGAGAGGCTTATCTGAA 1103
DB 234 GCGCCCTTACCAAGTGTCCCGGATCACCGCTCGACACCTGCGCCACGCGCGGAC 293
QY 1104 GCGCATTCGCTCGGTGAGCTTTCACGGGCGCGAGGACATGTCGAGAGAGACGC 1163
DB 294 GCGCTGGCTCGGAGACGAGACACCAACGACCCGAGGGCGCACGCGGATCATGGA 353
QY 1164 AAGCGCTGGGAGGTCAAGCGCGGTGAGCTTACGACGCGCAACTACACGCGCGGAGA 1223
DB 354 CGTGTGCGGCGCTGTGGAGACAGCTTCGACGACGAGCGCTTCGACAGACGCGCGCA 413
QY 1224 CAACATGACCAACCTCAACATGAAAGCTTCTGTATCACTTCAACGACGCGCTTGGC 1283
DB 414 GCGCTGTACTGCGCGCTGCGCCGCGCTCCACCACTTCACGACGAGGGAGGACATCGA 473
QY 1284 GTGCGCACTACGCGCAACGACCTCTCTGCG 1314
DB 474 CGTGGCCGCGCCCTCAACGATGCGCCGCGG 504

RESULT 9

T61016/c
ID T61016 standard; DNA: 30001 BP.
AC T61016;
AT 21-APR-1997 (first entry)
DE Total DNA sequence from cosmid clones LP(2)127 and LP(2)128.
KW Cosmid clone; LP(2)127; LP(2)128; chlorotetracycline;
RV biosynthetic pathway; recombinant; production; antibiotic;
KW heterologous host; Streptomyces lividans; ss.
OS Streptomyces aureofaciens.
PN US589385.A.
PD 31-DEC-1996.
PE 26-JUL-1990; 558039.
PR 26-JUL-1990; US-558039.
PR 26-JUL-1990; US-558040.
PR 15-JAN-1992; US-821419.
PR 15-JAN-1992; US-821419.
PR 22-SEP-1993; US-125468.
PA (AMCY) AMERICAN CYANAMID CO.
PI Fantini SE, Lotvin JA, Ryan MJ, Strathy N;
DR WPI: 97-076853/07.
PT DNA encoding tetracycline biosynthetic pathway proteins -
PT specifically from Streptomyces aureofaciens for expression in
PT heterologous hosts, specifically S. lividans
PS Claim 3; Columns 17-44; 39pp; English.
CC The present sequence is the total DNA sequence from cosmid clones
CC LP(2)127 and LP(2)128, which contains a Streptomyces aureofaciens
CC chlorotetracycline DNA that encodes the proteins of the entire
CC chlorotetracycline biosynthetic pathway. The biosynthetic gene,
CC which can be expressed in heterologous hosts, especially
CC S. lividans, may be useful in the production of antibiotics.
SQ Sequence 30001 BP; 4055 A; 10939 C; 11019 G; 3988 T;

Query Match 3.3%; Score 47.4; DB 1; Length 30001;
Best Local Similarity 45.9%; Pred. No. 0.036;
Matches 162; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 969 GCGCTTACACGAGGACCGCGGTTCGCGCCAGATGAGAAAGTGTTCGAGGACTA 1028
DB 8347 GCGACACGCGGACATGAGTGGCCAGCGGAGCGGTGGCCCGCGGCGTGAAGCT 8288
QY 1029 CTACGACTCATGCACTGCTCATATCCCAACCCAGTTTACGAGGACTTCCGACTGGA 1088
DB 8287 CTTCATTCGCGCCGACACGCGGAGGCGGTGCGGAGACGGTTCGCGCGGGGGA 8228
QY 1089 GGCCTTTATCTGAACGCAATTCGTCGGTGTGAGACTTTCACGGGCGCGGAGGATAT 1148
DB 8227 GGGCTGTAGGTGTGACGCGCTTCGCGGCGGAGCTCCGCGCGGAGCTGAGAGGCT 8168
QY 1149 GTCCGAGAGAGAACGCAACGCGCTGGGAGGTCAAGCGGCGGTGCTTCAGACGCCAATA 1208
DB 8167 GCTCCGGCGCGGCTGTGACCGCTTCGCGCGCATGCGCTGTGTCAACAACGCCGAGCG 8108
QY 1209 CCACGCGCGGAGAGACACATGACCAACCTCAACCATGAACTTCTGATCAATCCAA 1268
DB 8107 CAACGCGCGGTGGCGTCACTTCACCGCTCACCGAGAGCTGTGGAGAGACGTATGACAC 8048
QY 1269 AGCACCGCTTTCGCGCTGCGCACCTTACGCAACGACTTCTTCGATCCCA 1321
DB 8047 CAACCTCACGACGCTTTCGCGTACCCGCGAGGTCTCTCACACCGCGCGCA 7995

RESULT 10

X05110/c
ID X05110 standard; DNA: 30001 BP.
AC X05110;
AT 07-APR-1999 (first entry)
DE S. aureofaciens DNA in cosmid clones LP2-127 and LP2-128.
KW Chlorotetracycline; tetracycline; biosynthetic; actinomycete;
KW S. aureofaciens; S. ambofaciens; antibiotic resistance gene;
KW S. lividans; S. aureofaciens; cosmid clone; antibiotic; ss.
OS Streptomyces aureofaciens.
PN US5866410-A.
PD 02-FEB-1999.
PE 07-JUN-1995; 474933.
PR 22-SEP-1993; US-125468.
PR 26-JUL-1990; US-558039.
PR 26-JUL-1990; US-558040.
PR 15-JAN-1992; US-821419.
PR 15-JAN-1992; US-821419.
PR 07-JUN-1995; US-474933.
PA (AMCY) AMERICAN CYANAMID CO.
PI Fantini SE, Lotvin JA, Ryan MJ, Strathy N;
DR WPI: 99-141936/12.
PT Two plasmids for cloning the biosynthetic pathways of tetracycline,
PT chlorotetracycline, and their derivatives - comprise an origin of
PT replication, cohesive end sites and optionally, an antibiotic
PT resistance gene
PS Disclosure; Fig 4A-L; 39pp; English.
CC The invention relates to combination of two plasmids for cloning the
CC genes encoding the biosynthetic production pathways of chlorotetracycline,
CC tetracycline, or their derivatives. One plasmid contains an origin of
CC replication, an actinomycete (such as Streptomyces lividans,
CC S. aureofaciens, or S. ambofaciens) active antibiotic resistance gene and
CC three or more tandem cohesive end sites. The second plasmid contains an
CC actinomycete active origin of replication and three or more tandem
CC cohesive end sites. The invention can be used to clone a large amount of
CC genetic material for the heterologous production of the antibiotics,
CC chlorotetracycline, tetracycline and their analogues. A cluster of genes
CC can be inserted into the plasmid allowing a biosynthetic pathway to be
CC transferred in its entirety to a heterologous host. The invention
CC specifically relates to cloning of the entire tetracycline and
CC chlorotetracycline from Streptomyces aureofaciens and its expression in a
CC heterologous host such as S. lividans. The present sequence represents a
CC S. aureofaciens DNA sequence from the cosmid clones designated LP2-127

CC and LP2-128.
SO Sequence 30001 BP; 4055 A; 10939 C; 11019 G; 3988 T;

Query Match 3.3%; Score 47.4; DB 1; Length 30001;
Best Local Similarity 45.9%; Pred. No. 0.036;
Matches 162; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 969 GGGCTTACACGAGCGGCGGCGGCTTCCGCCACATGAGAAATGTTCCAGAGCTA 1028
DB 8347 GGGCACCAGCGGCGCTCGGACTGCGGACCGGCGGCGGCGGCGGCTGAGCT 8288
QY 1029 CTACAGCTCATGACCTGCGCTCATATCCACCCAGTTGAGAGGAGTTCAGTACGA 1088
DB 8287 CTTCATGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 8228
QY 1089 GGGCTTATCTGAGACGCAATTCGCTCGGTGACCTTTCACGGGCGGCGGCGGCTAT 1148
DB 8227 GGGGCTGAGAGTGAGCGGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 8168
QY 1149 GTCCGAGAGACAGCAGCGGCTGGGAGGTGACAGCGGCGGCGGCTACGACCGCACTA 1208
DB 8167 GGTCCGGGCGGCGGCTGAGCGGCTGCGGCGGCGGCGGCGGCTGCTGCTCAACAACGCGGAGC 8108
QY 1209 CCAAGCGCGGCGGAGACATGACCAACCTCAACATGAAGCCTTCTGATCAACTCCA 1268
DB 8107 CAACGGGCGGCTGAGCGGCTGACCTCAACCTCAACGAGAGCTGTGGAGAGCTGATCGACAC 8048
QY 1269 AGCCAGCGGCTTGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1321
DB 8047 CAACCTTACACGAGGCTTTCGCGGCTGACCGGCGGAGGTCTCAACGACCGGCGGCA 7995

RESULT 11

ID 064201 standard; CDNA; 5392 BP.
AC 064201;
DT 18-NOV-1994 (first entry)
DE Sequence comprising the *snaA*, *snaB* and *snaC* gene clusters.
KW Antibiotic: streptogramin; *snaA*, *snaB*, *snaC*; biosynthesis; enzyme:
biosynthetic pathway; Streptomyces pristinaespiralis; ds.
OS Streptomyces pristinaespiralis.
PN FR2696189-A.
PD 01-APR-1994.
PF 25-SEP-1992; 011441.
PR 25-SEP-1992; FR-011441.
PA (RHON) RHONE POUJENC ROGER SA.
PI Blanc V, Blanche F, Crouzet J, Jacques N, Lacroix P;
PI Thibaut D, Zagorec M;
DR MPI; 94-128286/16.
PT DNA involved in streptogramin antibiotic biosynthesis - for
PT prodn. or bio-conversion of streptogramin(s) or prodn. of
PT streptogramin intermediates, derivs. or hybrid antibiotics
PS Disclosure; Page 44-47; 83pp; French.
CC This sequence comprises the *snaA*, *snaB* and *snaC* genes which are
CC involved in the biosynthesis of streptogramins, antibiotics active
CC against Gram-positive bacteria. The identification of the sequences
CC encoding the enzymes involved in the biosynthetic pathway means that
CC they can be isolated and manipulated. Mutant microorganisms in
CC which a step in the streptogramin biosynthetic pathway is blocked
CC can be cultured to produce streptogramin intermediates, which may
CC later be converted to streptogramin derivatives. Recombinant cells
CC may also be used for the bioconversion of streptogramins from one
CC form to another or for the production of hybrid antibiotics.
SQ Sequence 5392 BP; 811 A; 2161 C; 1671 G; 749 T;

Query Match 3.2%; Score 46.2; DB 1; Length 5392;
Best Local Similarity 46.2%; Pred. No. 0.047;
Matches 153; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 984 CGAGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1043

DB 2865 CGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2924
QY 1044 CCGCTCATATCCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1103
DB 2925 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2984
QY 1104 CGGCTTCCGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1163
DB 2985 CGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3044
QY 1164 AAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1223
DB 3045 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3104
QY 1224 CAACATGACCAACCTCAACATGAGGCTTCTGATCAACTCAAGGCGGCGGCGGCGG 1283
DB 3105 GGGCTGTGACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3164
QY 1284 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1314
DB 3165 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3195

RESULT 12

ID 012569 standard; DNA; 1209 BP.
AC 012569;
DT 24-SEP-1991 (first entry)
DE Protease2 open reading frame from *A. chrysogenum*.
KW Promoter; cephalosporin; alkaline protease; lambda gt; ss.
OS Acromonium chrysogenum ATCC 11550.
FH Key location/Qualifiers
FT signal_peptide 1..351
FT mat_peptide 352..1206
FT /*tag= a
FT /*tag= b
PN EP-436355-A.
PD 10-JUL-1991.
PF 20-DEC-1990; 313988.
PR 27-DEC-1989; JP-342113.
PR 20-JUL-1990; JP-193609.
PA (FUJ1) FUJISAWA PHARM KK.
PI Isogai T, Fukagawa M, Iwami M, Aramori I, Kojo H;
DR MPI; 91-202062/28.
DR P-PSDB; R12975.
PT Single step fermentative prodn. of 7-amino-3-cephem cpds. - by
PT culturing Acromonium chrysogenum transformed with vector encoding
PT enzyme involved in N-deacylation
PS Claim 15; Page 47; 86pp; English.
CC The vector for producing 7-aminocephem cpd. contains a DNA fragment
CC prepared by ligating at least one or more promoter(s) for *A.*
CC *chrysogenum* to gene(s) for enzyme(s) capable of converting the
CC cephalosporin cpd. to the corresp. 7-aminocephem cpd. to each other
CC in that order from the upstream side to the downstream side by
CC conventional methods. An appropriate selective marker,
CC autonomously replicable sequence, terminator, translation activating
CC sequence and so forth may be inserted into the vector at respective
CC desired sites.
CC The promoter may be the promoter for the *A. chrysogenum* alkaline
CC protease gene as represented here.
SQ Sequence 1209 BP; 234 A; 482 C; 299 G; 194 T;

Query Match 3.0%; Score 43; DB 1; Length 1209;
Best Local Similarity 52.5%; Pred. No. 0.19;
Matches 94; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 1143 CATATGTCGAGAGAGCAAGCGGCTGGGAGGTCAAGCGGCGGCGGCTGAGAGCGC 1202
DB 786 CACAACAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 845
QY 1203 CAATACCAAGCGGCGGCGGAGACACATGACCAACTCAACCATGAAGCGCTTCTGATCAA 1262

DB 846 CGATACGAGAACGCCGCACTACTCCCCGGCTGGCGCCGCAACGCATGACCGTGG 905
 QY 1263 CTCGAAAGCCACCGCTTGGCGCTGCCACCTTACGCGACGCTCTCTGATCCCA 1321
 DB 906 CTCGATCGCTCCAACTGGCGCGCTCCAGCTTACGAACTACGCGTCCGCTCGACA 964

RESULT 13

Q12567
 ID Q12567 standard; cDNA: 1513 BP.
 AC Q12567;
 DT 24-SEP-1991 (first entry)
 DE Alkaline protease cDNA from A. chrysogenum.
 KW Promoter; cephalosporin; alkaline protease; ss.
 OS Acetomoniium chrysogenum ATCC 11550.
 PN EP-436355-A.
 PD 10-JUL-1991.
 PF 20-DEC-1990: 313988.
 PR 27-DEC-1989: JP-342113.
 PR 20-JUL-1990: JP-193609.
 PA (FUT1) FUJISAWA PHARM KK.
 PI Isogai T, Fukagawa M, Iwami M, Aramori I, Kojo H;
 DR WPI: 91-202062/28.
 PT Single step fermentative prodn. of 7-amino-3-cephem cpds. - by
 PT culturing Acetomoniium chrysogenum transformed with vector encoding
 PT enzyme involved in N-deacylation
 PS Disclosure; Fig 1-5-(1-2); 86pp; English.
 CC The vector for producing 7-aminocephem cpd. contains a DNA fragment
 CC prepared by ligating at least one or more promoter(s) for A.
 CC chrysogenum to gene(s) for enzyme(s) capable of converting the
 CC cephalosporin cpd. to the corresp. 7-aminocephem cpd. to each other
 CC in that order from the upstream side to the downstream side by
 CC conventional methods. An appropriate selective marker,
 CC autonomously replication sequence, terminator, translation activating
 CC sequence and so forth may be inserted into the vector at respective
 CC desired sites.
 CC The promoter may be the promoter for the A. chrysogenum isopenicillin
 CC N synthetase gene, the beta-isopropyl malate dehydrogenase gene or
 CC the alkaline protease gene as represented here. The promoter may
 CC contain an enhancer sequence.
 SQ Sequence 1513 BP; 303 A; 565 C; 382 G; 233 T;

Query Match 3.0%; Score 43; DB 1; Length 1513;

Best Local Similarity 52.5%; Pred. No. 0.2; Mismatches 85; Indels 0; Gaps 0;

Matches 94; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 1143 CAGCATGTCCGAGAGAGCAAGCCGCTGGGAGGTCAGCCGCGTGGCTACGACGC 1202
 DB 875 CAGCAAGCGCGTACACACTGCTACCTCCGCGGCTCTCTCCGCTGTGGCCGCGCA 934
 QY 1203 CAACTACGACGCGCGGAGAGCAACATGACCACTCAACACTGAAGCTTCTGATCAA 1262
 DB 935 CGATACGAGCAAGCGCGCAACTACTCCCGCGCTGGCGCGCAAGCCATACACGCGG 994
 QY 1263 CTCGAAAGCCACCGCTTGGCGCTGCCACCTTACGCGCAAGGACTCTCTGATCCCA 1321
 DB 995 CTCGATCGCTCCAACTGGCGCGCTCCAGCTTACGAACTACGCGTCCGCTCGACA 1053

RESULT 14

Q12568
 ID Q12568 standard; DNA: 1546 BP.
 AC Q12568;
 DT 24-SEP-1991 (first entry)
 DE Protease2 and beta-galactosidase fusion region.
 KW Promoter; cephalosporin; alkaline protease; lambda gt10; ss.
 OS Acetomoniium chrysogenum ATCC 11550.
 PN EP-436355-A.
 PD 10-JUL-1991.
 PF 20-DEC-1990: 313988.
 PR 27-DEC-1989: JP-342113.
 PR 20-JUL-1990: JP-193609.
 PA (FUT1) FUJISAWA PHARM KK.
 PI Isogai T, Fukagawa M, Iwami M, Aramori I, Kojo H;
 DR WPI: 91-202062/28.
 PT Single step fermentative prodn. of 7-amino-3-cephem cpds. - by
 PT culturing Acetomoniium chrysogenum transformed with vector encoding
 PT enzyme involved in N-deacylation
 PS Disclosure; Fig 1-4-(1-3); 86pp; English.
 CC The vector for producing 7-aminocephem cpd. contains a DNA fragment
 CC prepared by ligating at least one or more promoter(s) for A.
 CC chrysogenum to gene(s) for enzyme(s) capable of converting the
 CC cephalosporin cpd. to the corresp. 7-aminocephem cpd. to each other

FT EP-436355-A.
 PN 10-JUL-1991.
 PD 20-DEC-1990: 313988.
 PR 27-DEC-1989: JP-342113.
 PR 20-JUL-1990: JP-193609.
 PA (FUT1) FUJISAWA PHARM KK.
 PI Isogai T, Fukagawa M, Iwami M, Aramori I, Kojo H;
 DR WPI: 91-202062/28.
 DR P-PSDB: R12974.
 PT Single step fermentative prodn. of 7-amino-3-cephem cpds. - by
 PT culturing Acetomoniium chrysogenum transformed with vector encoding
 PT enzyme involved in N-deacylation
 PS Disclosure; Fig 1-6-(1-2); 86pp; English.
 CC The vector for producing 7-aminocephem cpd. contains a DNA fragment
 CC prepared by ligating at least one or more promoter(s) for A.
 CC chrysogenum to gene(s) for enzyme(s) capable of converting the
 CC cephalosporin cpd. to the corresp. 7-aminocephem cpd. to each other
 CC in that order from the upstream side to the downstream side by
 CC conventional methods. An appropriate selective marker,
 CC autonomously replication sequence, terminator, translation activating
 CC sequence and so forth may be inserted into the vector at respective
 CC desired sites.
 CC The promoter may be the promoter for the A. chrysogenum alkaline
 CC protease gene. The alkaline protease cDNA was obtd. based on
 CC expression of a protein fused with lambda gdt11 beta-galactosidase
 CC followed by antibody screening. The two frames are just in phase.
 CC The cDNA is linked to lambda gdt11 via a 21 bp GC tailing portion.
 SQ Sequence 1546 BP; 309 A; 565 C; 406 G; 266 T;

Query Match 3.0%; Score 43; DB 1; Length 1546;

Best Local Similarity 52.5%; Pred. No. 0.2; Mismatches 85; Indels 0; Gaps 0;

Matches 94; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 1143 CAGCATGTCCGAGAGAGCAAGCCGCTGGGAGGTCAGCCGCGTGGCTACGACGC 1202
 DB 903 CAGCAAGCGCGTACACACTGCTACCTCCGCGGCTCTCTCCGCTGTGGCCGCGCA 962
 QY 1203 CAACTACGACGCGCGGAGAGCAACATGACCACTCAACACTGAAGCTTCTGATCAA 1262
 DB 963 CGATACGAGCAAGCGCGCAACTACTCCCGCGCTGGCGCGCAACGATACCGTGG 1022
 QY 1263 CTCGAAAGCCACCGCTTGGCGCTGCCACCTTACGCGCAAGGACTCTCTGATCCCA 1321
 DB 1023 CTCGATCGCTCCAACTGGCGCGCTCCAGCTTACGAACTACGCGTCCGCTCGACA 1081

RESULT 15

Q12566
 ID Q12566 standard; DNA: 3172 BP.
 AC Q12566;
 DT 24-SEP-1991 (first entry)
 DE Alkaline protease gene from A. chrysogenum.
 KW Promoter; cephalosporin; alkaline protease; ss.
 OS Acetomoniium chrysogenum ATCC 11550.
 PN EP-436355-A.
 PD 10-JUL-1991.
 PF 20-DEC-1990: 313988.
 PR 27-DEC-1989: JP-342113.
 PR 20-JUL-1990: JP-193609.
 PA (FUT1) FUJISAWA PHARM KK.
 PI Isogai T, Fukagawa M, Iwami M, Aramori I, Kojo H;
 DR WPI: 91-202062/28.
 PT Single step fermentative prodn. of 7-amino-3-cephem cpds. - by
 PT culturing Acetomoniium chrysogenum transformed with vector encoding
 PT enzyme involved in N-deacylation
 PS Disclosure; Fig 1-4-(1-3); 86pp; English.
 CC The vector for producing 7-aminocephem cpd. contains a DNA fragment
 CC prepared by ligating at least one or more promoter(s) for A.
 CC chrysogenum to gene(s) for enzyme(s) capable of converting the
 CC cephalosporin cpd. to the corresp. 7-aminocephem cpd. to each other

Thu May 11 10:02:52 2000

us-09-080-127-1_copy_46_1488.rng

Page 10

CC in that order from the upstream side to the downstream side by
CC conventional methods. An appropriate selective marker,
CC autonomously replication sequence, terminator, translation activating
CC sequence and so forth may be inserted into the vector at respective
CC desired sites.
CC The promoter may be the promoter for the *A. chrysogenum* isopenicillin
CC N synthetase gene, the beta-isopropyl malate dehydrogenase gene or
CC the alkaline protease gene as represented here. The promoter may
CC contain an enhancer sequence.
SS Sequence 3172 BP; 654 A; 1103 C; 743 G; 669 T;

Query Match	3.0%	Score 43;	DB 1;	Length 3172;
Best Local Similarity	52.5%;	Pred. No. 0.23;		
Matches 94;	Conservative 0;	Mismatches 85;	Indels 0;	Gaps 0;

[illegible]

Search completed: May 10, 2000, 22:28:17
Job time: 2576 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2000, 21:43:19 ; Search time 63.05 seconds
(without alignments)
2905.969 Million cell updates/sec

Title: US-09-080-127-1_COPY_46_1488
Perfect score: 1443
Sequence: 1 GGGAGGGCGCTTGTTCGCC.....GGCATTCGAAAGTGAGGCA 1443

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 226296 seqs, 63486255 residues

Total number of hits satisfying chosen parameters: 452392

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/prodata/2/lna/5A_COMB.seq:*
2: /cgn2_6/prodata/2/lna/5B_COMB.seq:*
3: /cgn2_6/prodata/2/lna/5C_COMB.seq:*
4: /cgn2_6/prodata/2/lna/5D_COMB.seq:*
5: /cgn2_6/prodata/2/lna/6_COMB.seq:*
6: /cgn2_6/prodata/2/lna/PCFUS_COMB.seq:*
7: /cgn2_6/prodata/2/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	47.4	3.3	30001	1 US-08-125-468-1	Sequence 1, Appl
C 2	47.4	3.3	30001	3 US-08-474-933-1	Sequence 1, Appl
C 3	46.2	3.2	833	3 US-08-403-852D-3	Sequence 3, Appl
4	46.2	3.2	5392	3 US-08-403-852D-1	Sequence 1, Appl
5	43	3.0	474	3 US-08-403-852D-14	Sequence 14, Appl
6	43	3.0	1209	1 US-08-314-309A-5	Sequence 5, Appl
7	43	3.0	1513	1 US-08-314-309A-2	Sequence 2, Appl
8	43	3.0	1546	1 US-08-314-309A-3	Sequence 3, Appl
9	43	3.0	3172	1 US-08-314-309A-1	Sequence 1, Appl
10	43	3.0	5163	5 US-08-700-651-1	Sequence 1, Appl
11	43	3.0	5318	5 US-08-700-651-2	Sequence 2, Appl
12	42.6	3.0	459	4 US-08-387-942C-9	Sequence 9, Appl
C 13	40.8	2.8	13987	3 US-08-804-227C-13	Sequence 13, Appl
C 14	40.8	2.8	44377	3 US-08-804-227C-7	Sequence 7, Appl
C 15	40.8	2.8	44377	4 US-08-804-198-1	Sequence 1, Appl
16	40	2.8	1808	1 US-08-173-508-1	Sequence 1, Appl
17	40	2.8	1808	3 US-08-263-310-1	Sequence 1, Appl
C 18	39.8	2.8	1931	4 US-09-130-114-2	Sequence 4, Appl
19	39.6	2.7	1461	4 US-08-809-763-4	Sequence 4, Appl
20	39.6	2.7	12588	4 US-08-387-942C-1	Sequence 1, Appl
21	39.2	2.7	2064	1 US-08-343-428-1	Sequence 1, Appl
22	39	2.7	1229	2 US-08-440-861-1	Sequence 1, Appl
23	39	2.7	1229	2 US-08-433-854-1	Sequence 1, Appl
24	39	2.7	1229	2 US-08-174-745A-1	Sequence 1, Appl
25	39	2.7	1229	3 US-08-195-947-1	Sequence 1, Appl
26	39	2.7	1229	3 US-08-433-885-1	Sequence 1, Appl
27	39	2.7	1229	4 US-08-433-908B-1	Sequence 1, Appl

C 28	38.6	2.7	640	3 US-08-835-099A-16	Sequence 16, Appl
C 29	38.6	2.7	804	3 US-08-835-099A-10	Sequence 10, Appl
30	38.6	2.7	1437	5 US-08-724-814-15	Sequence 15, Appl
31	38.2	2.6	44377	4 US-08-804-227C-7	Sequence 7, Appl
32	38.2	2.6	44377	3 US-08-804-198-1	Sequence 1, Appl
33	38	2.6	1531	2 US-08-449-986-1	Sequence 1, Appl
34	38	2.6	1531	3 US-08-756-855-1	Sequence 1, Appl
35	38	2.6	28958	1 US-08-258-261B-6	Sequence 6, Appl
36	38	2.6	28958	1 US-08-456-837-6	Sequence 6, Appl
37	38	2.6	28958	1 US-08-457-342-6	Sequence 6, Appl
38	38	2.6	28958	1 US-08-457-646A-6	Sequence 6, Appl
39	38	2.6	28958	2 US-08-458-076A-6	Sequence 6, Appl
40	38	2.6	28958	2 US-08-764-232A-4	Sequence 4, Appl
41	38	2.6	28958	2 US-08-457-335A-6	Sequence 6, Appl
42	38	2.6	28958	2 US-08-729-214-6	Sequence 6, Appl
43	38	2.6	49377	2 US-08-764-233A-1	Sequence 1, Appl
44	37.8	2.6	2532	1 US-07-671-376C-4	Sequence 4, Appl
45	37.6	2.6	11219	2 US-07-642-734C-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-125-468-1/C
Sequence 1, Application US/08125468
Patent No. 5589385
GENERAL INFORMATION:
APPLICANT: Ryan, Michael J.
APPLICANT: Lotvin, Jason A.
APPLICANT: Strathy, Nancy A.
APPLICANT: Fanti, Susan E.
TITLE OF INVENTION: Cloning of the biosynthetic pathway for
TITLE OF INVENTION: Chlorotetracycline and tetracycline formation and cosmid
TITLE OF INVENTION: useful therein
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: USA
ZIP: 07470
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA: 31,145
APPLICATION NUMBER: US/08/125,468
FILING DATE: 22-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Teevos, Estelle J
REGISTRATION NUMBER: 31,145
REFERENCE/DOCKET NUMBER: 31,255-02
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201)831-3241
TELEFAX: (201)831-3305
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 30001 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-125-468-1
Query Match 3.3%; Score 47.4; DB 1; Length 30001;
Best Local Similarity 45.9%; Pred. No. 0.0084;
Matches 162; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

Qy	969	GGCCTTAAACAGACGGACCGCGGGCTTCCGCCAGATGAAAGAACTTTTGAGAGACTA	1028
Db	8347	GGGACACAGGGGCAATGGAGACTGGCCAGCCGGCAGCGGCTGTGCCCGGGCGGGGCTAAGCT	8288
Qy	1029	CTACGACTCCATGCAGCTGCTCTCATATCCCAACCCAGTTTGAAGAGACTTCCGACTACGA	1088
Db	8287	CTTATCTGCGCCCGCCACCGCGAGGGCGTGTGGCGAGAGAGTGTGCCCGCTCGGGGGGA	8228
Qy	1089	GGCCTTATCCTGAACGGCATTCCTGTCGGTGAATCTTACAGGGGCGGAGGGCGATCAT	1148
Db	8227	GGGCGCTGAAGTGAACGGCCTTCGGGGCGCAGTCCGCTCGCGCGGAGTGAAGAGGCT	8168
Qy	1149	GTCGGAAGAAACCAACCGCGCTGGGAGGTCAAGCGGCGTGGCCTTAGAGACCCAATA	1208
Db	8157	CGTCCGGGCGCGGTGAGACGGCTTCGGCGGGATCGCGCTGCTGTATACAAAGCCGAGCG	8108
Qy	1209	CCACGCCGCGGGAGACAAACATGACCAACCTCAACCATGAAAGCCTTCTGATCAACTCCA	1268
Db	8107	CAAGGGGAGTGAGCTCACTCCACCCCTCAACCGACCGAGCGAGCTGTGGACAGAGTATGACAC	8048
Qy	1269	AGGCACGCGCTTGCGGTGGCGACACTAAGCGCAACGACACTCTCTGATATCCCA	1321
Db	8047	CAACCTACCAAGGCTTTCGCGCTCAACCGCGAGAGTCTCTACCAACCGCGGGCA	7995

RESULT 2
US-08-474-933-1/c
; Sequence 1, Application US/08474933
; Patent No. 5866410

1 GENERAL INFORMATION:
 2 APPLICANT: Ryan, Michael J.
 3 APPLICANT: Lotvin, Jason A.
 4 APPLICANT: Strathby, Nancy
 5 APPLICANT: Pantini, Susan E.
 6 TITLE OF INVENTION: Cloning of the biosynthetic pathway for
 7 TITLE OF INVENTION: chlorotetracycline and tetracycline formation and cosmid
 8 TITLE OF INVENTION: useful therein
 9 NUMBER OF SEQUENCES: 1
 10 CORRESPONDENCE ADDRESS:
 11 ADDRESSEE: American Cyanamid Company
 12 STREET: One Cyanamid Plaza
 13

1
 2 CALL: Wayne
 3 STATE: New Jersey
 4 COUNTRY: USA
 5 ZIP: 07470
 6
 7 COMPUTER READABLE FORM:
 8 MEDIUM TYPE: floppy disk
 9 COMPUTER: IBM PC compatible
 10 OPERATING SYSTEM: PC-DOS/MS-DOS
 11 SOFTWARE: PatentIn Release #1.0, Version #1.25
 12 CURRENT APPLICATION DATA:
 13 APPLICATION NUMBER: US/08/474,933
 14 FILING DATE:
 15

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/125,468
FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Tsevdos, Estelle J
REGISTRATION NUMBER: 31,145
REFERENCE/DOCKET NUMBER: 3,255-020
TELECOMMUNICATION INFORMATION:

```

: TELEPHONE: (201)831-3241
:
: TELEFAX: (201)831-3305
:
: INFORMATION FOR SEO ID NO: 1:
:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 30001 base pairs
:   MAPPING: 1000000000

```

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
US-08-474-933-1

```

Query Match	3.3%;	Score 47.4;	DB 3;	Length 30001;
Best Local Similarity	45.9%;	Pred. No. 0.0084;		
Matches 162; Conservative	0;	Mismatches 191;	Indels 0;	Gaps 0;

Qy	969	GGCGTTTCAACCAAGACGGGACCGGGCGGTTCCGCCAGATCGAAGAACTTTCGAGACTA	1028
Db	8347	GGCGAACCGACGGGCAATCGGACTGGCGACGGCGGCGCGGTGGCCCGGGCTGAGCCT	8288
Qy	1029	CTAGACATCCATCGACTGGCTGATAATATCCCAACCGATTGACGAGACTTCGACTAGCA	1088
Db	8287	CTTATCTGGCGCCCGCACCCGCGACGGGCGTGGCGGAGACGGTGGCCCGCTCGCGGCGCA	8228
Qy	1089	GGCGTTTATCCGTGAACGGGCAATTCGTCGGGTGACATCTTCACGAGGCGCGAGGGCATCAT	1148
Db	8227	GGGCGCTGAGAGTGGACGCGCCTCGCGGCCACATCTCCCTCGCGCGGAGCTGACGAGCCT	8168
Qy	1149	GTCGGAAGAGAACGCAAGCGGCTGTGGGAGGTCAAGCGGGCGTGGCCCTACGAGCGCAACTA	1208
Db	8167	CGTCCGGGCGGCGGTGGACGCTTTCGGCGCGAGTGGCGCTGTGTGTCACACAGCGCGGAGC	8108
Qy	1209	CCAGCGCGCGGGGAGACAACTGACACCAACTCAACCATGAAAGCCTTCTATTAATCTCCA	1268
Db	8107	CAACGGGGGTGGCGCTCACTCCACCTTCACCGGACGAGCTGTGGCAGAGCATATGACAC	8048
Qy	1269	AGGCACGCGCTTGGCGCTGCACACTTACGCCAAGACACTCTCGTATCCCA	1321
Db	8047	CNACCTCAACAGGCTCTTCCGCGCTACCGCGAGAGTCTCTACACACCGCGGCGCA	7995

RESULT 3
US-08-403-852D-3
; Sequence 3, Application US/08403852D
; Patent No. 5891695
GENERAL INFORMATION:

1 GENERAL INFORMATION :
 2 APPLICANT: Blanc, Veronique
 3 APPLICANT: Blanche, Francis
 4 APPLICANT: Crouzet, Joel
 5 APPLICANT: Jacques, Nathalie
 6 APPLICANT: Lacroix, Patricia
 7 APPLICANT: Thibaut, Denis
 8 APPLICANT: Zagorec, Monique
 9 APPLICANT: Debussche, Laurent
 10 APPLICANT: De Creely-Lagard, Valerie
 11 TITLE OF INVENTION: Polypeptides Involved In The
 12 TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
 13 TITLE OF INVENTION: Coding For These Polypeptides And Their Use
 14 NUMBER OF SEQUENCES: 43
 15 CORRESPONDENCE ADDRESS:

ADDRESSSE: Flinnegan, Henderson, Farbow, Garrett & Dunne
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004-2915

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

APPLICANT NUMBER: US/08/403,852D
FILING DATE: 10-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93/11441

1 APPLICATION NUMBER: FR 9411441
2 FILING DATE: 25-SEP-1992
3 ATTORNEY/AGENT INFORMATION:
4 NAME: Meyers, Kenneth J.
5 REGISTRATION NUMBER: 25,146
6 REFERENCE/DOCKET NUMBER: 03806.0054-000000

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 833 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: S.pristinaespiralis
FEATURE:
NAME/KEY: CDS
LOCATION: 1..833
US-08-403-852D-3

Query Match 3.2%; Score 46.2; DB 3; Length 833;
Best Local Similarity 46.2%; Pred. No. 0.0049;
Matches 153; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 984 CGGACCGCGCGGTCCGCCAGATCGAAGAACTGTCGAGGACTACTACGACTGCATCGA 1043
DB 174 CGGCTGGCGCGCTCCGTCACGACATCGGCTATCACCGCCCGCTCCCGGCGACCA 233
QY 1044 CTTGCTCATATCCCAACCCAGTTTGACGAGCTTCGACTACGAGGCTTATCTGAA 1103
DB 234 GGGCCCTACACAGCTGTCCGATCACCCTGCTCGACCACTCCGCCACGGCCGAC 293
QY 1104 CGGCTTCGCTCGGTGACTTTCACGGGCGCCGAGGGCATCATGTCCGAGAAAGACGC 1163
DB 294 CGGCTGGCTCGGACGACGACACCGACCGCCGAGGCGCCGACCGGCAACTCATCGA 353
QY 1164 AAGCCGTGGGGAGGTGAAGCCGGCTGAGACGCAACTACACCGCGGGGAGA 1223
DB 354 CGTGTGCTGGCGGCTGTGGGACAGCTTCGACGACGACCGCTTCTGTCAGACCGCGCA 413
QY 1224 CAACATGACCAACTCAACCATGAAGCTTCTGATCAACTCAAGCCACCGCTTCGC 1283
DB 414 CGGCTGTACTGCGGGGTGCGCCGCTCCACCACTGACACCA3GGCAGCACTTGA 473
QY 1284 CGTGGCACTAGCGCAACGACTCTCTCG 1314
DB 474 CGTGGCGGCGCCCTCAACGTCGCGCCCG 504

RESULT 4
US-08-403-852D-1
Sequence 1, Application US/08403852D
Patent No. 5891695
GENERAL INFORMATION:
APPLICANT: Blanc, Veronique
APPLICANT: Blanc, Veronique
APPLICANT: Crouzet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibaut, Denis
APPLICANT: Zagorec, Monique
APPLICANT: Debussche, Laurence
APPLICANT: De Crey-Lagard, Valerie
TITLE OF INVENTION: Polypeptides Involved In The
TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
TITLE OF INVENTION: Coding For These Polypeptides And Their Use
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farbow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA

ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,852D
FILING DATE: 10-May-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PC/FR 93/00923
FILING DATE: 25-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806.0054-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5392 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: S.pristinaespiralis
US-08-403-852D-1

Query Match 3.2%; Score 46.2; DB 3; Length 5392;
Best Local Similarity 46.2%; Pred. No. 0.0094;
Matches 153; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 984 CGGACCGCGCGGTCCGCCAGATCGAAGAACTGTCGAGGACTACTACGACTGCATCGA 1043
DB 2865 CGGCTGGCGCGCTCCGTCACGACATCGGCTATCACCGCCCGCTCCCGGCGACCA 2924
QY 1044 CTTGCTCATATCCCAACCCAGTTTGACGAGCTTCGACTACGAGGCTTATCTGAA 1103
DB 2925 GGGCCCTACACAGCTGTCCGATCACCCTGCTCGACCACTCCGCCACCGCGCAC 2984
QY 1104 CGGCTTCGCTCGGTGACTTTCACGGGCGCCGAGGGCATCATGTCCGAGAAAGACGC 1163
DB 2985 CGGCTGGCTGGCGGCTGTGGGACAGCTTCGACGACGACCGCCGAGGCGCGCACTTGA 3044
QY 1164 AAGCCGTGGGGAGGTGAAGCCGGCTGAGACGCAACTACACCGCGGGGAGA 1223
DB 3045 CGTGTGCTGGCGGCTGTGGGACAGCTTCGACGACGAGCGCTTGTGTCACGACCGCGCGCA 3104
QY 1224 CAACATGACCAACTCAACCATGAAGCTTCTGATCAACTCAAGCCACCGCTTCGC 1283
DB 3105 CGGCTGTACTGCGGGGTGCGCCGCTCCACCACTGACACCA3GGCAGCACTTGA 3164
QY 1284 CGTGGCACTAGCGCAACGACTCTCTCG 1314
DB 3165 CGTGGCGGCGCCCTCAACGTCGCGCCCG 3195

RESULT 5
US-08-403-852D-14
Sequence 14, Application US/08403852D
Patent No. 5891695
GENERAL INFORMATION:
APPLICANT: Blanc, Veronique
APPLICANT: Blanc, Veronique
APPLICANT: Crouzet, Joel

```

1  APPLICANT: Jacques, Nathalie
2  APPLICANT: Lacroix, Patricia
3  APPLICANT: Thibaut, Denis
4  APPLICANT: Zagorec, Monique
5  APPLICANT: Debussche, Laurent
6  APPLICANT: De Crecy-Lagard, Valerie
7  TITLE OF INVENTION: Polypeptides Involved In The
8  TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
9  TITLE OF INVENTION: Coding For These Polypeptides And Their Use
10 NUMBER OF SEQUENCES: 43
11
12 CORRESPONDENCE ADDRESS:
13 ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
14 STREET: 1300 I Street, N.W., Suite 700
15 CITY: Washington
16 STATE: D.C.
17
18 COUNTRY: USA
19 ZIP: 20005-3315
20
21 COMPUTER READABLE FORM:
22 MEDIUM TYPE: Floppy disk
23
24 COMPUTER: IBM PC compatible
25 OPERATING SYSTEM: PC-DOS/MS-DOS
26 SOFTWARE: Patent Release #1.0, Version #1.30
27
28 CURRENT APPLICATION DATA:
29 APPLICATION NUMBER: US/08/403,852D
30 FILING DATE: 10-MAY-1995
31
32 PRIOR APPLICATION DATA:
33 APPLICATION NUMBER: PCT/FR 93/00923
34 FILING DATE: 25-SEP-1993
35
36 PRIOR APPLICATION DATA:
37 APPLICATION NUMBER: FR 92/11441
38 FILING DATE: 25-SEP-1992
39
40 ATTORNEY/AGENT INFORMATION:
41 NAME: Meyers, Kenneth J.
42 REGISTRATION NUMBER: 25,146
43 REFERENCE/DOCKET NUMBER: 03806.0054-00000
44
45 TELECOMMUNICATION INFORMATION:
46 TELEPHONE: (202) 408-4000
47 TELEFAX: (202) 408-4400
48
49 INFORMATION FOR SEQ ID NO: 14:
50 SEQUENCE CHARACTERISTICS:
51 LENGTH: 474 base pairs
52 TYPE: nucleic acid
53 STRANDEDNESS: double
54 TOPOLOGY: linear
55
56 MOLECULE TYPE: cDNA
57 HYPOTHETICAL: NO
58 ANTI-SENSE: NO
59
60 ORIGINAL SOURCE:
61 ORGANISM: S.pristinaespiralis
62 FEATURE:
63 NAME/KEY: CDS
64 LOCATION: 1..474
65 OTHER INFORMATION: /product= "Partie du gene Smbd"
66
67 US-08-403-852D-14

```

[illegible]

D _b	184	GCCCTGCTACACC	GGGCAGCGCGGCCG	CACCGACATCCCCGTGGCAGCCCCCTCCCGGC	243
Q _y	1255	CTGATCACTCCAA	AGCCACCGCCTTCG	CGCTGCCACCTACGCCAAGCATCTCTCTCG	1314
D _b	244	CGCACCGACGAC	---GCCCTCGAGAC	ACTGCTGGGTTTGTGTAACAACCTCGTCTG	300
Q _y	1315	ATCCCAACGAAGA	TATCACATCTCTCTTG	GCACGAGAGCCCGCACATGCGACATTG	1374
D _b	301	CGCACCGACACT	CGCGGACGCCACCTT	CCGGAAC TGTCGACAGCGTGC	360
Q _y	1375	GCG	1377		
D _b	361	GAC	363		

```

US-08-314-309A-5
US-08-314-309A-5
Sequence 5, Application US/08314309A
Patent No. 5677141
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: FUKAGAWA, MASAO
APPLICANT: IWAMI, MORITA
APPLICANT: ARAMORI, ICHIRO
APPLICANT: KOJO, HITOSHI
TITLE OF INVENTION: PROCESS FOR PRODUCING 7-AMINOCEPHEM
TITLE OF INVENTION: COMPOUND OR SALTS THEREOF
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSER: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSER: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/314,309A
FILING DATE: 30-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/631,906
FILING DATE: 21-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5677141man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-863-0 CONT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1209 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1206
US-08-314-309A-5

```

Query Match	3.0%	Score	43	DB	1	Length	1209
Best Local Similarity	52.5%	Pred. NO.	0.036				
Matches	94	Conservative	0	Mismatches	85	Indels	0
						Gaps	0

Db 963 CGATAACCGAACGCCCACTACTCCCCCGCCTCGGCCGCAACGCCATCACCCTCG 1022

Dy 1263 CTCCTAAAGCCACC GGCTTCGGCGTGGCCACCTTAGGCCAAGCACCCTCTCCTTGATCCCCA 1321
||||| ||||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1023 CTCCTATGCGCTCCAACGTGGGCGCTGCAGCTTCAGCAACTACGGCTCCGCTGCTGCACA 1081

RESULT 9

```

: US-08-314-309A-1
: Sequence 1, Application US/08314309A
: Patent No. 5677141
:
: GENERAL INFORMATION:
: APPLICANT: ISOGAI, TAKAO
: APPLICANT: FUKAGAWA, MASAO
: APPLICANT: IWAMI, MORITA
: APPLICANT: KAMORI, ICHIRO
: APPLICANT: KORO, HITOSHI
: TITLE OF INVENTION: PROCESS FOR PRODUCING 7-AMINOCEPHEM
: TITLE OF INVENTION: COMPOUND OR SALTS THEREOF
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESS: OBION, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
: ADDRESSEE: P. C.
: STREET: 1755 S. Jefferson Davis Highway, Suite 400
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/314,309A
: FILING DATE: 30-SEP-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/631,906
: FILING DATE: 21-DEC-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: Obion, No. 5677141man F.
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 18-863-0 CONT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 413-3000
: TELEFAX: (703) 413-2220
: TELEX: 248855 OPAT UR
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3172 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: unknown
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
:
: US-08-314-309A-1

```

Query Match	3.0%:	Score 43:	DB 1:	Length 3172:
Best Local Similarity	52.5%:	Pred. No. 0.05:		
Matches	94:	Conservative	0:	Mismatches 85: Indels 0: Gaps 0:
QY 1143	CATCATGTCCGAGAGAGACGCAAGCCGCTGGGSGAGTCAAGCCGCGTGGCTACGACGC	1202		
DB 1907	CACACAGCGCGTCAACACTGCTCTACTCCGGGGGCTCTCTCCGTCGTGAGCGCGCGCA	1966		
QY 1203	CACCTACCAAGCGCCGGGAGACACACATATCCCAACTCAACCTGAAGCTTCTGATCAA	1262		
DB 1967	CGATACACAGAACCCCGGCACTACTCCCGCGCTCGGCGCGCAAGCGCATACCGGTGG	2026		
QY 1263	CTCCAAAGCCACCGCTTTCGCGCTGCGCCACCTAGGCGCAAGCGACTCTCTGATGCCCA	1321		
DB 2027	CTCATGCGCTTCAACTATGGGCGCCGCTCCAGCTTACGCAACATATAGGCTCTCGTCTGACA	2085		

RESULT 10
 US-08-700-651-1
 Sequence 1, Application US/08700651B
 Patent No. 6015882
 GENERAL INFORMATION:
 APPLICANT: PETERSEN, CAROLYN
 APPLICANT: LEECH, JAMES
 APPLICANT: NELSON, RICHARD, C.
 APPLICANT: GUY, JIRI
 TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
 TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
 TITLE OF INVENTION: INFECTIONS
 FILE REFERENCE: 480.19-4(HV)
 CURRENT APPLICATION NUMBER: US/08/700,651B
 CURRENT FILING DATE: 1997-08-14
 EARLIER APPLICATION NUMBER: 08/415,751
 EARLIER FILING DATE: 1995-04-03
 NUMBER OF SEQ ID NOS: 15
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 1
 LENGTH: 5163
 TYPE: DNA
 ORGANISM: Cryptosporidium parvum
 US-08-700-651-1

Query Match	Similarity	Score	DB	Length
Best Local Similarity	42.9%	Pred. No. 0	05	5163
Matches	214	Conservative	0	Mismatches 285; Indels 0; Gaps 0;
QY	849	CAGCACTACTAGCTGCTCCACTGTAATGCGCAGCGAGCTGACAGATCGACTGTACCT	908	
Db	591	caactactactactactactactacgacaacaacaacaacaacagcaacaacaactactacta	650	
QY	909	GAACTTGCAKATGATGCGCTCACTTACTATACGCCCTCATGATCTATATAGCGGTATGATC	968	
Db	651	caactactactaccactactactactagacaacaacaacaacaacaacaacaacaacaaca	710	
QY	969	GGCGTTCACACAGAGCGGAGCGGGCGGTTCGGCCAGATGAGAAACTGTGAGGACTA	1028	
Db	711	caacaacaacaacaacaacaacaacgactactactactactactactactactactacta	770	
QY	1029	CTACGACTCATCGACCTGCGCTCATATCCGCCACCAGTTGACGAGAGCTTCGACTACGA	1088	
Db	771	ctactactacaacacacaacaactactacaacccaactactacaacacaacaactacaac	830	
QY	1089	GGCCTTATCTTGAAAGGCAATTCGTCGGGTGACTCTTCACGGGGCGCGAGGGCATAT	1148	
Db	831	caacaacaacaacaacaactacaacccaactataacacacaacaactacaacacaaca	890	
QY	1149	GTCGAGAGAGACGACGAGCGGCTGGGGAGTGCANAGCGCGGCTGGCTTACGACGGCAACTA	1208	
Db	891	ccacaacccaacaacacaactaccagaagaaaccaacaacaacaactacacaacaaca	950	
QY	1209	CCAGCGCGCGGAGACACATGACCAACTTACCATTAAGAACTTCTCGATCAACTCCAA	1268	
Db	951	caataaacaactactactactacaacacacaacaabaacaacaacaacaacaactacta	1010	
QY	1269	AGGCACGCGCTTGGCGGTGCGCACTTACGCGACGACACTTCTCGATCCCCAAGGAA	1328	
Db	1011	agaacaacaactactactactactaccacaacaacaacaacaactactactaccacaa	1070	
QY	1329	TACCAKATCTCTTGCAC	1347	
Db	1071	caacaacaactactactac	1089	

RESULT 11

US-08-700-651-2
; Sequence 2, Application US/08700651B
; Patent No. 6015882

GENERAL INFORMATION:
APPLICANT: PETERSEN, CAROLYN
APPLICANT: LEECH, JAMES
APPLICANT: NELSON, RICHARD, C.
APPLICANT: GUY, JIRI
TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAs AND RNAs
TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
TITLE OF INVENTION: INFECTIONS
FILE REFERENCE: 480.19-4(hv)
CURRENT APPLICATION NUMBER: US/08/700.651B
CURRENT FILING DATE: 1997-08-14
EARLIER APPLICATION NUMBER: 08/415,751
EARLIER FILING DATE: 1995-04-03
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2
LENGTH: 5318
TYPE: DNA
ORGANISM: Cryptosporidium parvum
US-08-700-651-2

Query Match 3.0%; Score 43; DB 5; Length 5318;
Best Local Similarity 42.9%; Pred. No. 0.06;
Matches 214; Conservative 0; Mismatches 285; Indels 0; Gaps 0;

QY 849 CAGCACTACTACTCTCCCTCCATGATGCCACGAGCTGAACAAGATCCGACTGTACT 908
DB 591 CAACTACTACTACTACTACTACTGACAAACAACAACAACAACAACAACAACAACA 650
QY 909 GAACCTTGACATGATCGCTCACTACTACTACTACTACTACTACTACTACTACTACT 968
DB 651 CAACCTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 710
QY 969 GGCCTTCAACCAAGCGGACCGCGGCTCCGCGAGATCGAAGAAATGTTGAGAGACTA 1028
DB 711 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 770
QY 1029 CTAGACTCTCATGACCTGCTCATATCCCAACCAAGTTTACGAGAGCTTCCGACTACGA 1088
DB 771 CTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 830
QY 1089 GGCCTTATCTGTAACGGGATTCGCTCGGTGACTCTTCAAGGGGCGCGAGGATCAT 1148
DB 831 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 890
QY 1149 GTCCGAAGAGACGCAACCGCTGGGAGGTCAAGCCGCGTGGCTTACGAGCCCACTA 1208
DB 891 CCACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 950
QY 1209 CCAGCGCGCGGAGACAGACATGACCACTCAACCATTAAGCTTCTGATCACTACTCA 1268
DB 951 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1010
QY 1269 AGCCACGCGCTTCCGCGGCGGACCTAGCCCAAGCAAGCTTCTGATCCCAAGCGAA 1328
DB 1011 AGAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1070
QY 1329 TACCAATCTCTCTTGCAC 1347
DB 1071 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1089

RESULT 12
US-08-387-942C-9
Sequence 9, Application US/08387942C
Patent No. 5939289
GENERAL INFORMATION:
APPLICANT: ERTSVAG, HELGA
APPLICANT: VALLA, SVEIN
APPLICANT: SKJAK-BRAEK, GUDMUND
APPLICANT: LARSEN, BJORN
TITLE OF INVENTION: DNA COMPOUNDS COMPRISING SEQUENCES

TITLE OF INVENTION: ENCODING MANNURONAN C-5-EPIMERASE
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
STREET: P.O. BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/387,942C
FILING DATE: 09-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY JR, GERALD M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1809-106P
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
INFORMATION FOR SRO ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-387-942C-9

Query Match 3.0%; Score 42.6; DB 4; Length 459;
Best Local Similarity 51.9%; Pred. No. 0.033;
Matches 96; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 1132 GCGCGGAGGAGCATATGTCGGAAGAGCAAGCCGTGGGAGGTCAAGCCGCGTG 1191
DB 91 GCGCGGAGGAGCATATGTCGGAAGAGCAAGCCGTGGGAGGTCAAGCCGCGTG 150
QY 1192 GCGTACGAGCGCACTACACCGCGCGGAGAGACATGACCAACCTCAACCAAGAGCC 1251
DB 151 GCGCGGAGGAGCATATGTCGGAAGAGCAAGCCGTGGGAGGTCAAGCCGCGTG 210
QY 1252 TTCCTGATCACTCAAGGACCGCGCTTGGCGGCACTACGCACTAGCCCAAGAGCTCTCC 1311
DB 211 GACAGCTTCAAGGACGATGATCAGGACTTGAGCGCAGGAGGAGCGGATGAGCTGTCC 270
QY 1312 TCGAT 1316
DB 271 GCGCT 275

RESULT 13
US-08-804-227C-13/C
Sequence 13, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kuhstoss, Stuart A.
APPLICANT: Kosteck, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA

ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 13987 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 350..13987
US-08-804-227C-13

Query Match 2.8%; Score 40.8; DB 3; Length 13987;
Best Local Similarity 47.6%; Pred. No. 0.3;
Matches 120; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 1116 CGGTGACCTTCACAGGGGCGGAGGCGCATGTCCGAGAGAACGCAAGCCGCTGGGG 1175
DB 10366 CGGCCATTTCAGCGCGGACGACACCCGCTCCCAACCCCTCAACCCCGCGCAC 10307
QY 1176 AGGTCAAGCGGCGTGGCTTACGACGCGCAACTACGACGCGGCGGAGACAATGACCAA 1235
DB 10306 CACGCACTTCAGAGCGCCGACGACGACATCCCGCCACCGGAGAACCCACCATCA 10247
QY 1236 CCTCAACCATGAAGCCTTCGTATCACTCAACTCAAGGACCGGCTTGGCGGCACCTA 1295
DB 10246 CCCACCGCGCGACACGACGACGACGACGACGACGACGACGACGACGACGACGAC 10187
QY 1296 CGCCAGCAAGCCTCTCTCTGATCCCAAGGAAATACCAATCTCTTTCACGAGAGAC 1355
DB 10186 CGCAGCAGCAGCCTCCCGACGAAATGACCAACACGACGACGACGACGACGACG 10127
QY 1356 CGCAGCATGCG 1367
DB 10126 CTCGAAGCGCG 10115

RESULT 14

US-08-804-227C-7/c
Sequence 7, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: DeHoff, Bradley S.
APPLICANT: Kustoss, Stuart A.
APPLICANT: Rostock, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 44377 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 350..14002
FEATURE:
NAME/KEY: CDS
LOCATION: 14046..20036
FEATURE:
NAME/KEY: CDS
LOCATION: 20110..31284
FEATURE:
NAME/KEY: CDS
LOCATION: 31329..36071
FEATURE:
NAME/KEY: CDS
LOCATION: 36155..41830
US-08-804-227C-7

Query Match 2.8%; Score 40.8; DB 3; Length 44377;
Best Local Similarity 47.6%; Pred. No. 0.45;
Matches 120; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 1116 CGGTGACCTTCACAGGGGCGGAGGCGCATGTCCGAGAGAACGCAAGCCGCTGGGG 1175
DB 16316 CGGCCATTTCAGCGCGGACGACGACGACATCCCGCCAAACCCCTCAACACCCCGCAC 16257
QY 1176 AGGTCAAGCGGCGTGGCTTACGACGCGCAACTACGACGCGCGGAGAGACAATGACCAA 1235
DB 16256 CACGCACTTCAGAGCGCCGACGACGACGACGACGACGACGACGACGACGACGAC 16197
QY 1236 CCTCAACCATGAAGCCTTCGTATCACTCAACTCAAGGACCGGCTTGGCGGCACCTA 1295
DB 16196 CCCACCGCGCGACACGACGACGACGACGACGACGACGACGACGACGACGACGAC 16137
QY 1296 CGCCAGCAAGCCTCTCTCTGATCCCAAGGAAATACCAATCTCTTTCACGAGAGAC 1355
DB 16136 CGCAGCAGCAGCCTCCCGACGAAATGACCAACGACGACGACGACGACGACGACG 16077
QY 1356 CGCAGCATGCG 1367
DB 16076 CTCGAAGCGCG 16065

RESULT 15

US-08-804-198-1/c
Sequence 1, Application US/08804198
Patent No. 5945320
GENERAL INFORMATION:
APPLICANT: Burgett, Stanley G.
APPLICANT: Kustoss, Stuart A.
APPLICANT: Rao, Nagara R.
APPLICANT: Richardson, Mark A.

Thu May 11 10:02:55 2000

us-09-080-127-1_copy_46_1488.rni

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2000, 21:01:18 ; Search time 572.26 Seconds

(without alignments)
10220.531 Million cell updates/sec

Title: US-09-080-127-1_COPY_46_1488

Perfect score: 1443
Sequence: 1 GGGAGGGCGCTGTTCGCC.....GGCATTCTCAATCGAGCA 1443

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 202611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 08
Listing first 45 summaries

Database : EST:*

1:	em_est1:*
2:	em_est2:*
3:	em_est3:*
4:	em_est4:*
5:	em_est5:*
6:	em_est6:*
7:	em_est7:*
8:	em_est8:*
9:	em_est9:*
10:	em_est10:*
11:	em_est11:*
12:	em_est12:*
13:	em_est13:*
14:	em_est14:*
15:	em_est15:*
16:	em_est16:*
17:	em_est17:*
18:	em_est18:*
19:	em_est19:*
20:	gb_est1:*
21:	gb_est2:*
22:	gb_est3:*
23:	gb_est4:*
24:	gb_est5:*
25:	gb_est6:*
26:	gb_est7:*
27:	gb_est8:*
28:	gb_est9:*
29:	gb_est10:*
30:	gb_est11:*
31:	gb_est12:*
32:	gb_est13:*
33:	gb_est14:*
34:	gb_est15:*
35:	gb_est16:*
36:	gb_est17:*
37:	gb_est18:*
38:	gb_est19:*
39:	gb_est20:*
40:	gb_est21:*
41:	gb_est22:*
42:	gb_est23:*
43:	gb_est24:*
44:	gb_est25:*

45:	gb_est26:*
46:	gb_est27:*
47:	gb_est28:*
48:	gb_est29:*
49:	gb_est30:*
50:	gb_est31:*
51:	gb_est32:*
52:	em_est20:*
53:	em_est21:*
54:	em_est22:*
55:	em_est23:*
56:	em_est24:*
57:	em_est25:*
58:	em_est26:*
59:	gb_est33:*
60:	gb_est34:*
61:	gb_est35:*
62:	gb_est36:*
63:	gb_est37:*
64:	gb_est38:*
65:	em_est27:*
66:	em_est28:*
67:	em_est29:*
68:	em_est30:*
69:	gb_est39:*
70:	gb_est40:*
71:	gb_est41:*
72:	gb_est42:*
73:	gb_est43:*
74:	gb_est44:*
75:	em_est31:*
76:	em_est32:*
77:	em_est33:*
78:	em_est34:*
79:	gb_est45:*
80:	gb_est46:*
81:	gb_est47:*
82:	gb_est48:*
83:	gb_est49:*
84:	gb_est50:*
85:	gb_est51:*
86:	em_est35:*
87:	em_est36:*
88:	em_est37:*
89:	em_est38:*
90:	gb_est52:*
91:	gb_est53:*
92:	gb_est54:*
93:	gb_est55:*
94:	gb_est56:*
95:	em_est39:*
96:	em_est40:*
97:	em_est41:*
98:	em_est42:*
99:	em_est43:*
100:	em_est44:*
101:	em_est45:*
102:	gb_est57:*
103:	gb_est58:*
104:	em_est46:*
105:	gb_est59:*
106:	gb_est60:*
107:	gb_est61:*
108:	gb_est62:*
109:	gb_est63:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	131.8	9.1	492	43	AI213187	AI213187 y8d01a1.f
2	119.8	8.3	625	74	AJ273485	AJ273485 AJ273485
3	113.8	7.9	604	74	AJ274177	AJ274177 AJ274177
4	110.2	7.6	565	74	AJ273341	AJ273341 AJ273341
5	109.8	7.6	618	74	AJ274066	AJ274066 AJ274066
6	102.2	7.1	635	74	AJ272856	AJ272856 AJ272856
7	94.2	6.5	600	74	AJ272843	AJ272843 AJ272843
8	93.4	6.5	480	74	AJ273602	AJ273602 AJ273602
9	63.2	4.4	291	38	AA784895	AA784895 y8d01a1.f
10	63.2	4.4	314	43	AI213186	AI213186 y8d01a1.f
11	57.6	4.0	341	38	AA784896	AA784896 y8d01a1.f
12	56	3.9	606	74	AJ273199	AJ273199 AJ273199
13	54.4	3.8	589	74	AJ272994	AJ272994 AJ272994
14	49.8	3.5	925	82	CNS0091P	AL053013 Drosophila
15	47.2	3.3	599	79	AM289155	AM289155 707010D07
16	47.2	3.3	939	82	CNS00CNG	AL059400 Drosophila
17	47	3.3	427	79	AM288814	AM288814 707010D06
18	46.4	3.2	608	81	AM400334	AM400334 707059C03
19	46.2	3.2	234	37	AA712486	AA712486 32201.fam
20	46.2	3.2	892	85	AO687294	AO687294 nbxb00746
21	45.8	3.2	1101	82	CNS00K2	AL077673 Drosophila
22	45.6	3.2	566	108	AO577805	AO577805 nbxb0091P
23	45.6	3.2	571	105	AO326688	AO326688 nbxb0038A
24	45.6	3.2	633	105	AO3280802	AO3280802 nbxb0037F
25	45.6	3.2	676	45	AI389106	AI389106 GH20192.5
26	44.8	3.1	884	82	CNS00600	AL065923 Drosophila
27	44.2	3.0	783	41	AI069045	AI069045 m9ae0005b
28	43.2	3.0	1101	83	CNS0172M	AI107512 Drosophila
29	43	3.0	330	38	AA5454278	AA5454278 97MJ0121
30	42.8	3.0	401	90	AO794246	AO794246 nbxb0052I
31	42.8	3.0	511	62	AO762298	AO762298 AU076298
32	42.8	3.0	689	62	AU076299	AU076299 AU076299
33	42.8	3.0	922	82	CNS0073W	AL066784 Drosophila
34	42.8	3.0	1101	83	CNS00175X	AL108460 Drosophila
35	42.6	3.0	895	82	CNS00071A	AL066286 Drosophila
36	42.6	3.0	925	82	CNS0006RY	AL065849 Drosophila
37	42	2.9	574	61	AI861086	AI861086 603011G12
38	42	2.9	611	38	AA754477	AA754477 97SN1825
39	41.8	2.9	587	107	AO449445	AO449445 m9xb0023H
40	41.6	2.9	783	41	AI069358	AI069358 m9ae0006d
41	41.6	2.9	787	41	AI069371	AI069371 m9ae0006d
42	41.6	2.9	910	41	AI069241	AI069241 m9ae0006a
43	41.6	2.9	926	70	AM155575	AM155575 m9ae00037C
44	41.4	2.9	322	20	D22141	D22141 R1C010405A
45	41.4	2.9	362	40	C91669	C91669 C91669 Rice

ALIGNMENTS

RESULT 1
LOCUS AI213187 492 bp mRNA EST 19-OCT-1998
DEFINITION y8d01a1.f1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans cDNA clone y8d01a1 5', mRNA sequence.

ACCESSION AI213187
VERSION AI213187.1 GI:3775129
KEYWORDS EST.
SOURCE Emericella nidulans.
ORGANISM Emericella nidulans.
Eukaryota; Fungi; Ascomycota; Eurotiales; Trichocomaceae; Emericella.

REFERENCE 1 (bases 1 to 492)
AUTHORS Kupfer,D., Gray,J., Hausner,J., Lai,H., Martin,W., Aramayo,R., Prade,R. and Roe,B.
TITLE An Aspergillus nidulans EST Database
JOURNAL Unpublished (1998)
COMMENT Other ESTs: y8d01a1.f1
Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry

Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center
Seq primer: T3
High quality sequence stop: 431.
Location/Qualifiers
1..492
/organism="Emericella nidulans"
/strain="FGSC A26"
/db_xref="taxon:5072"
/clone="y8d01a1"
/clone_lib="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library"
/tissue_type="vegetative mycelia, asexual structures"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; 5' end of cDNA cloned into XhoI site of pBluescript 3' end of cDNA cloned into XhoI site of pBluescript"

FEATURES
source
BASE COUNT 118 a 117 c 132 g 125 t
ORIGIN

Query Match 9.1%; Score 131.8; DB 43; Length 492;
Best Local Similarity 55.9%; Pred. No. 4.3e-19;
Matches 278; Conservative 0; Mismatches 207; Indels 12; Gaps 1;

QY 640 AACCTTGTGGCGACAGACGAGGCGGCGATCCGAACAGTGTGCGCGTGGTGGCCAC 699
DB 8 AACATTATGGCGACAGACGAGGCGGACCTGAAACCTGTAATGCTCGAGGCTAC 67
QY 700 ACGGCTAGTCTGAGCGCGCGCTGTCTCAACGACATGCTCGGCGATATTAGCAAC 759
DB 68 AGTGCACGCTGGCGGAGGAGCTGTATCAATGACGACGCTGTGGACCTGACCTT 127
QY 760 TTGGCATTTGCCAAGCGCTCAGCAGTACTCGTCAAGATGCGTGGCGTTCCTTC 819
DB 128 CTGGAAGTGGCGACCACTTATGCAATATGACGATGATGATGATGATGATGATG 187
QY 820 TGGACAGGAGAGAGTTCGTCTGTGGGAGCACTACTAGCTTCCATGATGATG 879
DB 188 TGGCGGCTGGAGGAGGAGGAGGCTGTGGCTCCGACTATATGATGCTCAGGAA 247
QY 880 ACCGAGCTGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 939
DB 248 GAGGAGACCTGAGATTCGTCTTATGATGATGATGATGATGATGATGATGATG 307
QY 940 GCCCTCATGATCTATGACGATGATGATGATGATGATGATGATGATGATGATGATG 999
DB 308 GCTATCAAGTCTACAA-----TCTACGAATAGGTTAACCCTGTGTGATCC 355
QY 1000 GCCCAGATCGAAGACCTGTTCGAGGACTACTAGCACTCCATGACCTGCTCATATCCC 1059
DB 356 GAAGAGCTACGATCTTATACGAGGTTTACACTTCTCATGCGCTCAACTACATAC 415
QY 1060 ACCAGTTTACGAGAGCTTCCGACTAGAGGCTTTATCCGAAAGGCGATTCGCGGT 1119
DB 416 ATTCGTTGATGGAGAGAGTATGATGATGATGATGATGATGATGATGATGATG 475
QY 1120 GGACTCTTCACGGCGC 1136
DB 476 GGTATCGTACTGGTGC 492

RESULT 2
LOCUS AJ273485 625 bp mRNA EST 29-DEC-1999
DEFINITION AJ273485 Metarhizium anisopliae ARSEF 2575 Metarhizium anisopliae cDNA clone Ma#1058, mRNA sequence.
ACCESSION AJ273485
VERSION AJ273485.1 GI:6432858

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	COMMENT
RESULT 4 AJ273341	AJ273341	565 bp	MRNA	EST	29-DEC-1999					
	AJ273341 Metarhizium anisopliae ARSEF 2575 Metarhizium anisopliae									
	CDNA clone Ma#874, mRNA sequence.									
	AJ273341									
	AJ273341.1	GI:6432713								
	EST.									
	Metarhizium anisopliae.									
	Metarhizium anisopliae.									
	Eukaryota; Fungi; Ascomycota; anamorphic Ascomycota; Metarhizium.									
	1 (bases 1 to 565)									
	Screen,S.E., Mathur,P. and St. Leger,R.J.									
	EST analysis of the insect pathogenic fungus Metarhizium anisopliae									
	Unpublished (1999)									
	On Dec 20, 1995 this sequence version replaced gi:1134366.									
	Contact: Screen SE									
	Entomology									
	University of Maryland									
	4112 Plant Sciences Building, College Park, MD 20742, USA.									
	Location/Qualifiers									
	1..565									
	/organism="Metarhizium anisopliae"									
	/strain="ARSEF 2575"									
	/db_xref="taxon:5530"									
	/clone="Ma#874"									
	/clone_id="Metarhizium anisopliae ARSEF 2575"									
	/note="Vector: Unizap. Metarhizium anisopliae was grown on									
	insect cuticle for 24 hours. A cDNA library was									
	constructed in the unidirectional Lambda vector, Unizap"									
	BASE COUNT	122 a	172 c	173 g	98 t					
	ORIGIN									
	Query Match	7.6%	Score 110.2:	DB 74:	Length 565:					
	Best Local Similarity	52.4%	Pred. No. 2.2e-14:							
	Matches 291:	Conservative 0:	Mismatches 258:	Indels 6:	Gaps 2:					
Oy	392	GAGATGCCCCGCTGCGGACAGTGGGTCTGCTGCGCAACCAAGGCGCGCTTCA	451							
Db	1	GAGCGTGGCCCTTTCGCCATGATGAGCTCTGCGCAAGGACCGGACCAAGCGCTCA	60							
Oy	452	TTGTCTATACAACTGTGGCGGATCCATGCGCGGACACCTT--GGCGCGCGCAAGTG	508							
Db	61	TCTTTACAAACACAGCGCCGCGGACAGAACTACAGACCGCTGCGCGGCAAAACG	120							
Oy	509	ATAAGGAAACCGATTCGGCCATTTGCGGATACGTTGGAGATGGCCAAAGCTGATCA	568							
Db	121	TGCGCAAGCTGATTCGGCTCCGGCGGCTGTGGCCCTTAGAGACGCTGAACCTGGCGCGC	180							
Oy	569	AGCTTGTGAGCGCTGATCGGATCTGATCTGGGTGATAGTAAGCAGGAAC	628							
Db	181	GGTGTGCGCGGCGGACCAAGCTCACGCTCAACCTGTGTGTGATCCGCTTCGAAACCC	240							
Oy	629	GTCACAGCTATACGTTGTGCGGACAGAGAGGCGCGGATCCGAACAAGCTGTCGCC	688							
Db	241	GCGAAGCTGGAGACATCACTCTCGGAGACCAAGCCGCGGACAAAGACATATGCTCATGC	300							
Oy	689	TGGGTGGCCACAGGACATCAGTCGAGGGGGGCGCTGGATCAAGAGATGGCTGGGGCA	748							
Db	301	TGGGCGCCATCTGGACAGCGTCTCCGCGGGGCCGCGGGCTCAAGCAGGCGAGGAA	360							
Oy	749	TTATTAGCAACTGTGTCATTGGCAAGGCGCTCACGACTACCG--TCAAGATGCGG	805							
Db	361	CCGCGCGCTTGTGTGAGATTATCGAGAGGTTATTCGATATAGCGGCAATCAAGATTAGG	420							

DB	481	CCAACCTCACCGAGCGGAGAGCCGACATCAAGTCTACTTCAACATGATGATG	540
QY	926	CCTCACCTACTAGC 940	
Db	541	GATCGCCCATCCG 555	
RESULT	5		
LOCUS	AJ274066	618 bp mRNA	EST 29-DEC-1999
DEFINITION	AJ274066	Metarhizium anisopliae ARSEF 2575 Metarhizium anisopliae cDNA clone Xa#1689, mRNA sequence.	
ACCESSION	AJ274066		
VERSION	AJ274066.1	GI:6433439	
KEYWORDS	EST.		
SOURCE	Metarhizium anisopliae.		
ORGANISM	Metarhizium anisopliae		
REFERENCE	1 (bases 1 to 618)		
AUTHORS	Screen, S.E., Mathur, P. and St. Leger, R.J.		
JOURNAL	Est. analysis of the insect pathogenic fungus Metarhizium anisopliae		
COMMENT	Unpublished (1999)		
	On Dec 20, 1995 this sequence version replaced gi:135580.		
	Contact: Screen SE		
FEATURES	Entomology		
	University of Maryland		
	4112 Plant Sciences Building, College Park, MD 20742, USA.		
	Location/Qualifiers		
	1..618		
	/organism="Metarhizium anisopliae"		
	/strain="ARSEF 2575"		
	/db_xref="taxon:5530"		
	/clone="Xa#1689"		
	/clone_id="Metarhizium anisopliae ARSEF 2575"		
	/note="Vector: Unizap; Metarhizium anisopliae was grown on insect cuticle for 24 hours. A cDNA library was constructed in the unidirectional lambda vector, Unizap"		
BASE COUNT	132 a	182 c	202 g 100 t 2 others
ORIGIN			
Query Match	7.6%;	Score 109.8;	DB 74; Length 618;
Best Local Similarity	54.1%;	Pred. No. 2.8e-14;	
Matches	269;	Conservative 0;	Mismatches 222; Indels 6; Gaps 2;
QY	358	GATGTCGAGGGCAGAGTCCCTGATCAAGCGTGAGAGTCCCGTTCGGCGACAGTCG	417
Db	77	GATGCCCGAGGAAAGGTTCCTGCTGAAGGAGAGGCTCGCTTTCGCGATGACG	136
QY	418	GTTCTCGTGCACAAAGCCAAAGCCGGCGCTTCGATGTCTATACAAATGTGGCCGATCC	477
Db	137	CTGCTGGCAAGAGCAGCCGAGCCAGGCGCTCATCTTCTCAACAAACACCCCGGCAAG	196
QY	478	ATGGGCGGACACCTT---GGCGGGCGCAGATATAAGGACCGTATTCGGCCATTGTC	534
Db	197	AACATGACACACCGCTACGCTGACAGCGCAAAACGTCCGCAAGCTGATTCGTCGGGCTC	256
QY	535	GGATACAGCTTGAAGATGAGCCAGAAAGCTGATCAAGCTTCTGAGAGCTGATCGATATCT	594
Db	257	GTTGCCCTAAGGAGAGCGGTGAAGACCTGGGCGCGGCTGTGCGGCGGACAAAGTCAAGC	316
QY	555	GTGATCTGTGTGGGTGATAGTAAGCAGAGAAACCGTACGATTAACGTTGTGCGCAG	654
Db	317	GTCACCCCTGTGTGTCGATGACCGTCTCCGAAACCCGCGAGAGTGAACATCATCTCGGAG	376
QY	655	ACGAAGGCGGCGGATCCGACAAACAGTGTGCGGCTGGGTGGGCGCACAGGATCAGTCAAG	714
Db	377	ACCAAGCGCGCGAGAACAAATGTGTCTACTGCTGGGCGCCCATCTGGAGAGCGTCTC	436
QY	715	GCAGGCGCTGTATCAAGCAGATGAGCTCGGCGCATATTATCAACTGTTGTCATTCGCAAA	774
Db	437	CCGGGCGCGCGGCGTCAAGCAGAGCGCAAGCGGAACCGCGCTTGTGTGGAATATTATGAG	496

LOCUS	AI213186	314 bp	mRNA	EST	19-Oct-1998
DEFINITION	y8d01a1.f1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans cDNA clone y8d01a1 3', mRNA sequence.				
ACCESSION	y8d01a1 3' mRNA sequence.				
VERSION	AI213186				
KEYWORDS	AI213186.1	GI:3775128	EST.		
SOURCE	Emericella nidulans.				
ORGANISM	Emericella nidulans. Eukaryota; Fungi; Ascomycota; Eurotiiales; Trichocomaceae; Emericella.				
REFERENCE	1 (bases 1 to 314) Kudfer,D., Gray,J., Hausner,J., Lal,H., Martin,W., Aramayo,R., Prade,R. and Roeb,B. An Aspergillus nidulans EST Database Unpublished (1998)				
TITLE	Other-ESTs: y8d01a1.r1				
JOURNAL	Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu Department of Chemistry and Biochemistry Advanced Center for Genome Technology, University of Oklahoma 620 Parrington Oval, Norman, OK 73019, USA Tel: 405 325 4912 Fax: 405 325 7762				
COMMENT	Email: broe@ou.edu We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center Seq primer: M13-20 High quality sequence stop: 274. Location/Qualifiers 1..314 /organism="Emericella nidulans" /strain="FGSC A26" /db_xref="taxon:5072" /clone="y8d01a1" /clone_lib="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library" /issue_type="vegetative mycelia, asexual structures" /note="Vector: pluescript SK-; Site.1: EcoRI; Site.2: XhoI; 5' end of cDNA cloned into EcoRI site of pluescript 3' end of cDNA cloned into XhoI site of pluescript"				
FEATURES	source				
BASE COUNT	73 a	90 c	72 g	79 t	
ORIGIN					

[illegible]

```

FEATURES             SOURCE
SOURCE
vegetative cDNA lambda zap library Emericella nidulans cDNA clone
93b06a1.5', mRNA sequence.
ACCESSION            AA/84896
VERSION              AA/84896.1
KEYWORDS             GI:2845064
SOURCE               EST.
ORGANISM             Emericella nidulans.
                    Emericella nidulans
                    Eukaryota; Fungi; Ascomycota; Eurotiales; Trichocomaceae;
                    Emericella.
REFERENCE            1 (bases 1 to 341)
AUTHORS             Kupfer,D., Gray,J., Hausner,J., Lai,H., Martin,W., Aramayo,R.,
                    Prade,R. and Roe,B.
TITLE               An Aspergillus nidulans EST Database
JOURNAL             Unpublished (1998)
COMMENT             On Jan 14, 1998 this sequence version replaced gi:1797443.
                    Other-ESTs: 93b06a1.f1
                    Contact: Bruce A. Roe, University of Oklahoma, broeou.edu
                    Department of Chemistry and Biochemistry
                    Advanced Center for Genome Technology, University of Oklahoma
                    620 Parrington Oval, Norman, OK 73019, USA
                    Tel: 405 325 4912
                    Fax: 405 325 7762
                    Email: broeou.edu
                    We anticipate the future release of the cDNA clones to the Fungal
                    Genetics Stock Center
                    Seq primer: SK
                    High quality sequence stop: 324.
                    Location/Qualifiers
                    1..341

```

BASE COUNT	78 a	96 c	86 g	81 t
ORIGIN				
Query Match	4.0%	Score 57.6:	DB 38;	Length 341;
Best Local Similarity	53.6%	Pred. NO. 0.005;		
Matches 120;	Conservative 0;	Mismatches 104;	Indels 0;	Gaps 0;
QY	271	ACCATGACCTACACTCCAGCGCTGAGGCTCACCCGCCGATGTAGCCGTGTCAGAACCTG	330	
DB	80	ACCCTCCACACAGAACAGAGAGCAGTAAACGGCTGCTAAATCCTGTGTGCAGACAA	139	
QY	331	GGATGCACGAGCGCGGATTCACCATCCGATCGGATCGAGGCGCAAGTGGCCCTGATCAAGCT	390	
DB	140	GGATGCGCAAGGGCGAGCTATCTCTTGAGCTCCCTGGGCGCAATTCCTTGATCAACGC	199	
QY	391	GGAGATATCCCGCTTCGGGAGCAAGTCGGTTCTGCTGCGCAAGCGCAAGCGCGGCTTG	450	
DB	200	GGTACCTCTCTTCCTTCGGTACCAATCAAGACTAGCTGGCAAGCAGAGCGCGAGTTGCCGC	259	
QY	451	ATTGCTATTAACAATGTGGCCGATTCATGGCGGCGCAACCTTGG	494	
DB	260	GTAGATATCAATTAACGAGCATGTGAAGTTAGCGGAACGTTGG	303	
RESULT 12				
LOCUS	AJ273199	606 bp	EST	29-DEC-1999
DEFINITION	AJ273199 Metarhizium anisopliae ARSEF 2575 Metarhizium anisopliae			
ACCESSION	AJ273199			
VERSION	AJ273199.1 GI:6432571			
KEYWORDS	EST.			
SOURCE	Metarhizium anisopliae.			

ORGANISM

Metarhizium anisopliae

Eukaryota; Fungi; Ascomycota; anamorphic Ascomycota; Metarhizium.

REFERENCE 1 (bases 1 to 606)

AUTHORS Screen, S.E., Mathur, P. and St. Leger, R.J.

TITLE EST analysis of the insect pathogenic fungus Metarhizium anisopliae

JOURNAL Unpublished (1999)

COMMENT On Dec 20, 1995 this sequence version replaced g1:1134833.

Contact: Screen SE

Entomology University of Maryland

4112 Plant Sciences Building, College Park, MD 20742, USA.

Location/Qualifiers

1. 606

/organism="Metarhizium anisopliae"

/strain="ARSEF 2575"

/db_xref="taxon:5530"

/clone="Ma#664"

/note="Vector: Unizap; Metarhizium anisopliae ARSEF 2575"

Insect cuticle for 24 hours. A cDNA library was

constructed in the unidirectional Lambda vector, Unizap"

BASE COUNT 136 a 176 c 172 g 121 t 1 others

ORIGIN

Query Match 3.9%; Score 56; DB 74; Length 606;

Best Local Similarity 55.0%; Pred. No. 0.014;

Matches 132; Conservative 0; Mismatches 105; Indels 3; Gaps 1;

QY 719 GCCCTGATATCAACGACGATGCGTCGATATAGCACTTGGTCAATGCGCAAGCGC 778

Db 1 GGGCCGGGGGTACAGACGACGACGACGACGACGACGACGACGACGACGACGACG 60

QY 779 TCAGCC---AGTACTCGCTCAAGAAATGCCGTCGCTTCTCTCTGACAGCAGAGAGT 835

Db 61 TGATTCGATATGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 120

QY 836 TCGGCTCTGCGGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 895

Db 121 GCGGCTTATCGATCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 180

QY 896 TCCGACTGTACTGACATGATGCGCTCACTAATGACGCGCTCATGATCTATG 955

Db 181 TCAAGTCTACTTCACTACGACATGATGATGATGATGATGATGATGATGATGATG 240

RESULT 13

AJ272994 589 bp mRNA EST 29-DEC-1999

LOCUS AJ272994 Metarhizium anisopliae ARSEF 2575 Metarhizium anisopliae

DEFINITION CDNA clone Ma#333, mRNA sequence.

ACCESSION AJ272994.1 GI:6432367

VERSION AJ272994.1 GI:6432367

KEYWORDS Metarhizium anisopliae.

SOURCE Metarhizium anisopliae

ORGANISM Eukaryota; Fungi; Ascomycota; anamorphic Ascomycota; Metarhizium.

REFERENCE 1 (bases 1 to 589)

AUTHORS Screen, S.E., Mathur, P. and St. Leger, R.J.

TITLE EST analysis of the insect pathogenic fungus Metarhizium anisopliae

JOURNAL Unpublished (1999)

COMMENT On Dec 20, 1995 this sequence version replaced g1:1133782.

Contact: Screen SE

Entomology University of Maryland

4112 Plant Sciences Building, College Park, MD 20742, USA.

Location/Qualifiers

1. 589

/organism="Metarhizium anisopliae"

/strain="ARSEF 2575"

/db_xref="taxon:5530"

/clone="Ma#333"

/clone_11b="Metarhizium anisopliae ARSEF 2575"

FEATURES

source

/note="Vector: Unizap; Metarhizium anisopliae was grown on insect cuticle for 24 hours. A cDNA library was constructed in the unidirectional Lambda vector, Unizap"

BASE COUNT 135 a 168 c 166 g 120 t

ORIGIN

Query Match 3.8%; Score 54.4; DB 74; Length 589;

Best Local Similarity 57.7%; Pred. No. 0.03; 71; Indels 0; Gaps 0;

Matches 97; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 788 ACTGCTCAAGAAATGCCGTCCTCTCTGACAGCAGAGAGAGAGAGAGAGAGAGAG 847

Db 14 ACGGATCAAGAAATGCGTTGCGTCGCGTGGTGGGAGAGAGAGAGAGAGAGAGAGAG 73

QY 848 GCACCACTACTACGTTCCATCTGAATGCCAGCGAGTGAACAGATCCAGTACC 907

Db 74 GATGTTGTTACTACGTTCCAACTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 133

QY 908 TGACTTCGACATGATGCGCTCACTACGAGCGCCATGATCTATG 955

Db 134 TCAACTACGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 181

RESULT 14

CNS0091P 925 bp DNA GSS 03-JUN-1999

LOCUS CNS0091P/c

DEFINITION Drosophila melanogaster genome survey sequence TERT end of BAC #

BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit

fly), genomic survey sequence.

AL053013

VERSION AL053013.1 GI:4934461

KEYWORDS fruit fly.

SOURCE Drosophila melanogaster

ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 925)

AUTHORS

TITLE Direct Submission

JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;

BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila genome project (BDGP).

The BDGP is constructing a physical map of the Drosophila

melanogaster genome using these BACs. For further information

please see <http://www.fruitfly.org> The BDGP Drosophila

melanogaster BAC library was prepared by Kazutoyo Osoegawa and

Aaron Mammeter at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPCI-98 and was constructed by partial

EcoRI digestion of Drosophila DNA provided by the BDGP from the

isogenic strain y2; cn bw sp, the same strain used for the BDGP's

P1 and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or

filters for hybridization from the BACPAC Resource Center can be

found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

Location/Qualifiers

1. 925

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/clone_11b="RPCI-98"

/clone="BACR19D16"

/note="end : TERT"

BASE COUNT 120 a 172 c 172 t 511 others

ORIGIN

Query Match 3.5%; Score 49.8; DB 82; Length 925;

Best Local Similarity 13.3%; Pred. No. 0.36;

Matches 48; Conservative 163; Mismatches 150; Indels 0; Gaps 0;

Thu May 11 10:02:59 2000

us-09-080-127-1_copy_46_1488.rst

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2000, 22:26:09 ; Search time 1003.22 Seconds
(without alignment):
-1445.777 Million cell updates/sec

Title: US-09-080-127-1
Perfect score: 1491
Sequence: 1 ATGAGTGCCTTTGTGGGC.....ATTCTCAAGTCGAGCATAG 1491

Scoring table: OLIGO-NUC
Gapop 60.0 , Gapext 60.0

Searched: 882769 seqs, -486395729 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl:.*
1: gb_bal:.*
2: gb_bal:.*
3: gb_bal:.*
4: gb_bal:.*
5: gb_bal:.*
6: gb_bal:.*
7: gb_bal:.*
8: gb_bal:.*
9: gb_bal:.*
10: gb_bal:.*
11: gb_bal:.*
12: gb_bal:.*
13: gb_bal:.*
14: gb_bal:.*
15: gb_bal:.*
16: gb_bal:.*
17: gb_bal:.*
18: gb_bal:.*
19: gb_bal:.*
20: gb_bal:.*
21: gb_bal:.*
22: gb_bal:.*
23: gb_bal:.*
24: gb_bal:.*
25: gb_bal:.*
26: gb_bal:.*
27: gb_bal:.*
28: gb_bal:.*
29: gb_bal:.*
30: gb_bal:.*
31: gb_bal:.*
32: gb_bal:.*
33: gb_bal:.*
34: gb_bal:.*
35: gb_bal:.*
36: gb_bal:.*
37: gb_bal:.*
38: gb_bal:.*
39: gb_bal:.*
40: gb_bal:.*
41: gb_bal:.*
42: gb_bal:.*
43: gb_bal:.*

44: gb_htg6:.*
45: gb_htg7:.*
46: em_htg1:.*
47: em_htg2:.*
48: em_htg3:.*
49: em_htg4:.*
50: gb_pl3:.*
51: gb_pl5:.*
52: gb_pl8:.*
53: gb_pl9:.*
54: gb_pl10:.*
55: gb_pl11:.*
56: gb_pl12:.*
57: gb_pl13:.*
58: gb_pl14:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	1.3	2440	1 PDETFAB	L14864 Paracoccus
2	20	1.3	7390	2 AFCHERRA	J05278 Ralstonia e
3	20	1.3	126105	45 AC011881	AC011881 Homo sapi
4	19	1.3	128	5 A94242	A94242 Sequence 7
5	19	1.3	1878	1 REJ001389	AJ001389 Rhizobium
6	19	1.3	1891	7 VARBNASQA	223132 V. alboarum
7	19	1.3	5026	4 GGU16848	U16848 Gallus gall
8	19	1.3	12486	2 U39714	U39714 Mycoplasma
9	19	1.3	85518	35 AC005449	AC005449 Drosophila
10	19	1.3	112460	11 HS1104E15	AL022312 Human DNA
11	19	1.3	145873	57 AC011325	AC011325 Homo sapi
12	19	1.3	154700	32 AP000803	AP000803 Homo sapi
13	19	1.3	157559	44 AC020344	AC020344 Drosophila
14	18	1.2	208	13 AF033522	AF033522 HyA regl
15	18	1.2	349	13 G51319	G51319 SHGC-82833
16	18	1.2	709	13 G53977	G53977 SHGC-89037
17	18	1.2	1118	12 AF009133	AF009133 Rattus no
18	18	1.2	1159	8 AF110783	AF110783 Volvox ca
19	18	1.2	1994	7 VDRNNSQA	Z23133 V. danilaela
20	18	1.2	2207	34 TC0A9146	AJ009146 Trypanoso
21	18	1.2	2217	34 TC0A9145	AJ009145 Trypanoso
22	18	1.2	2240	34 TC0223563	AJ223563 Trypanoso
23	18	1.2	2260	34 TC0A9144	AJ009144 Trypanoso
24	18	1.2	2457	9 HSCACRTR	X66403 H. sapiens m
25	18	1.2	3348	16 S62819	S62819 F2L-putativ
26	18	1.2	3353	12 MUSGCE	L41933 Mus musculu
27	18	1.2	3512	12 AF084482	AF084482 Mus muscu
28	18	1.2	3530	12 MM0011971	AJ011971 Mus muscu
29	18	1.2	3645	12 RATGCD	L37203 Rattus norv
30	18	1.2	4362	40 AF105999	AF105999 Homo sapi
31	18	1.2	5892	16 AB009614	AB009614 Leucania
32	18	1.2	6969	34 TC022319	U22319 Trypanosoma
33	18	1.2	6991	34 TC022315	U22315 Trypanosoma
34	18	1.2	7216	8 AF104926	AF104926 Verticill
35	18	1.2	8304	16 AF091736	AF091736 VESV-like
36	18	1.2	8598	35 DM035631	U35631 Drosophila
37	18	1.2	11273	1 AB023289	AB023289 Pseudomon
38	18	1.2	20303	41 AC006749	AC006749 Caenorhab
39	18	1.2	23195	2 AF035413	AF035413 Agrobacte
40	18	1.2	24916	34 CELK353	L15313 C. elegans
41	18	1.2	32269	42 AC015257	AC015257 Drosophila
42	18	1.2	33606	43 AC017861	AC017861 Drosophila
43	18	1.2	33779	1 SCGD3	AL096822 Streptomy
44	18	1.2	34193	35 CELT22H9	AF101315 Caenorhab
45	18	1.2	36300	1 MTC128	Z97050 Mycobacteri
46	18	1.2	37338	34 CELF32B5	AF003148 Caenorhab
47	18	1.2	37750	1 SC6D7	AL133213 Streptomy
48	18	1.2	38000	41 AC010079	AC010079 Leishmani

C 49	18	1.2	40544	1	SCHS	AL035636 Streptomy	122	17	1.1	1412	8	AF020567	AF020567 Juniperus
C 50	18	1.2	45111	43	AC017279	AC017279 Drosophill	123	17	1.1	1423	11	HSLRPS520	AF058416 Homo sapi
C 51	18	1.2	51430	43	AC017981	AC017981 Drosophill	124	17	1.1	1479	31	PIGNPA	M25547 Pig brain n
C 52	18	1.2	53450	1	MTV018	AL021899 Mycobacte	125	17	1.1	1481	16	AF131648	AF131648 Hyposoter
C 53	18	1.2	55782	44	AC019933	AC019933 Drosophill	126	17	1.1	1488	5	I32147	I32147 Sequence 4
C 54	18	1.2	57827	45	AC021204	AC021204 Homo sapi	127	17	1.1	1497	5	I32146	I32146 Sequence 3
C 55	18	1.2	65821	55	AC020826	AC020826 Mus muscu	128	17	1.1	1509	5	E05391	E05391 Bile acid-s
C 56	18	1.2	77737	45	AC005640	AC005640 *** SEQUE	129	17	1.1	1509	5	I32145	I32145 Sequence 1
C 57	18	1.2	78260	42	AC015439	AC015439 Drosophill	130	17	1.1	1583	16	D88474	D88474 Hepatitis C
C 58	18	1.2	78515	42	AC021928	AC021928 Homo sapi	131	17	1.1	1597	9	CAMOSC	X12449 African Gre
C 59	18	1.2	84485	35	AC004351	AC004351 Drosophill	132	17	1.1	1645	40	AF131778	AF131778 Homo sapi
C 60	18	1.2	89938	35	AC006459	AC006459 Drosophill	133	17	1.1	1672	35	U96158	U96158 Drosophilla
C 61	18	1.2	110000	32	CEY105C5-1	Continuation (2 of	134	17	1.1	1677	50	AF160511	AF160511 Yarrowia
C 62	18	1.2	110000	52	AC010277_4	Continuation (5 of	135	17	1.1	1686	1	SLI250495	AJ250495 Streptomy
C 63	18	1.2	116857	11	HSDJ655K7	AL050342 Human DNA	136	17	1.1	1719	1	CU156275	CU156275 Commomys s
C 64	18	1.2	129127	11	HSDJ177G6	AL078639 Human DNA	137	17	1.1	1720	16	AF156933	AF156933 Hyposoter
C 65	18	1.2	131188	11	HSJ1103B4	AL121998 Human DNA	138	17	1.1	1740	1	AB019238	AB019238 Setralia
C 66	18	1.2	133609	52	AC021416	AC021416 Homo sapi	139	17	1.1	1900	3	BOVPHOS	L18966 Bos laurus
C 67	18	1.2	135436	54	AC008496	AC008496 Homo sapi	140	17	1.1	1936	12	MUSP580TA	M58633 Mouse p58/g
C 68	18	1.2	137387	10	HSJ468K18	AL049844 Homo sapi	141	17	1.1	1977	4	AF031897	AF031897 Meleagris
C 69	18	1.2	150343	45	AC005973	AC005973 Homo sapi	142	17	1.1	2049	9	AB02474S02	AB024748 Homo sapi
C 70	18	1.2	153164	44	AC015913	AC015913 Homo sapi	143	17	1.1	2077	1	ENTAC3YI	M88012 Enterobacte
C 71	18	1.2	155661	42	AC010921	AC010921 Drosophill	144	17	1.1	2077	1	AB000508	AB000508 Comamonas
C 72	18	1.2	159421	33	AL136323	AL136323 Homo sapi	145	17	1.1	2113	5	A96072	A96072 Sequence 10
C 73	18	1.2	159763	43	AC010344	AC010344 Homo sapi	146	17	1.1	2151	5	A96074	A96074 Sequence 10
C 74	18	1.2	163000	52	AC021385	AC021385 Homo sapi	147	17	1.1	2151	5	A96078	A96078 Sequence 11
C 75	18	1.2	174320	44	AC016987	AC016987 Homo sapi	148	17	1.1	2253	5	E07167	E07167 DNA encodin
C 76	18	1.2	174732	42	AC012160	AC012160 Drosophill	149	17	1.1	2256	5	E07166	E07166 DNA encodin
C 77	18	1.2	175582	33	AL138817	AL138817 Homo sapi	150	17	1.1	2262	16	LDG38895	U38895 Lymantiria d
C 78	18	1.2	182615	41	AC009595	AC009595 Homo sapi	151	17	1.1	2267	5	E07165	E07165 DNA encodin
C 79	18	1.2	182655	52	AC010740	AC010740 Homo sapi	152	17	1.1	2272	7	YSCAMY	E07165 Saccharomyc
C 80	18	1.2	184635	41	AC007989	AC007989 Homo sapi	153	17	1.1	2342	1	DMFUSG	X68014 D.mobilis F
C 81	18	1.2	185652	40	AC007312	AC007312 Homo sapi	154	17	1.1	2466	12	AF151363	AF151363 Mus muscu
C 82	18	1.2	192448	52	AC013413	AC013413 Homo sapi	155	17	1.1	2526	2	SMACHITIN	L41660 Serratia ma
C 83	18	1.2	192717	33	AL138935	AL138935 Homo sapi	156	17	1.1	2532	12	SCYRR286W	Z36155 S.cerevisia
C 84	18	1.2	198839	33	AL138955	AL138955 Homo sapi	157	17	1.1	2534	7	RNRNABGPI	X6406 R.norvegicu
C 85	18	1.2	200392	53	AC010976	AC010976 Homo sapi	158	17	1.1	2535	5	E07168	E07168 DNA encodin
C 86	18	1.2	203199	43	AC005721	AC005721 Drosophill	159	17	1.1	2619	4	X1A224945	AJ224945 Xenopus l
C 87	18	1.2	207684	45	AC018039	AC018039 Drosophill	160	17	1.1	2627	7	YSCSEPIA	M63577 S.cerevisia
C 88	18	1.2	210964	45	AC009075	AC009075 Homo sapi	161	17	1.1	2745	11	HSN252060	AJ252060 Homo sapi
C 89	18	1.2	217240	41	AC010898	AC010898 Homo sapi	162	17	1.1	2853	10	HSNCAD	X54315 Human mRNA
C 90	18	1.2	218494	32	AL133333	AL133333 Homo sapi	163	17	1.1	2868	12	RNALPARE	X79881 R.norvegicu
C 91	18	1.2	224040	37	AC007618	AC007618 Homo sapi	164	17	1.1	3052	7	SCYBR198C	Z36067 S.cerevisia
C 92	18	1.2	236735	32	AL136104	AL136104 Homo sapi	165	17	1.1	3077	12	RNT37142	U37142 Rattus norv
C 93	18	1.2	286205	45	AC009196	AC009196 Homo sapi	166	17	1.1	3153	12	MBREVEGEN	X87096 M.musculu
C 94	18	1.2	314838	34	CEY105C5A	AL117193 Caenorhab	167	17	1.1	3160	10	RRI005294	U05294 Rhodospirill
C 95	17	1.1	100	10	HSNCADX9A	Z27427 H.sapiens N	168	17	1.1	3161	12	MUSCDPK	L37092 Mus musculu
C 96	17	1.1	224	40	HSPOLE24	AF127960 Homo sapi	169	17	1.1	3448	9	HUMNCADH	M34064 Human N-cad
C 97	17	1.1	264	13	AU049737	AU049737 Rattus no	170	17	1.1	3451	2	U01158	X57548 Human mRNA
C 98	17	1.1	283	9	HS72ABF	M62937 H.sapiens C	171	17	1.1	3671	2	AF076952	AF076952 Homo sapi
C 99	17	1.1	291	1	STMG1AB	AF126336 Drosophill	172	17	1.1	3862	40	AF076952	AF076952 Homo sapi
C 100	17	1.1	438	35	AF162636	AF162636 Drosophill	173	17	1.1	3936	51	AF156539	AF156539 Homo sapi
C 101	17	1.1	796	2	AF181078	AF181078 Rhodobact	174	17	1.1	3986	2	U97022	U97022
C 102	17	1.1	939	40	AF081053	AF081053 Propithec	175	17	1.1	4120	1	SEHGTABCP	U97022 Fervidobact
C 103	17	1.1	939	40	AF081055	AF081055 Varecia v	176	17	1.1	4130	10	SEHGTABCP	X92556 S.erythraea
C 104	17	1.1	939	40	AF081056	AF081056 Varecia v	177	17	1.1	4132	10	S42303	S42303 N-cadherin
C 105	17	1.1	939	40	AF081057	AF081057 Hapalemur	178	17	1.1	4473	31	SEPCCBBCP	X92557 S.erythraea
C 106	17	1.1	939	40	AF081058	AF081058 Lemur cat	179	17	1.1	4761	35	AF047881	X92557 S.erythraea
C 107	17	1.1	939	40	AF081059	AF081059 Eulemur f	180	17	1.1	4768	35	AF047880	AF047881 Caenorhab
C 108	17	1.1	939	40	AF081060	AF081060 Eulemur f	181	17	1.1	4797	2	AB023185	AF047880 Caenorhab
C 109	17	1.1	939	40	AF081061	AF081061 Eulemur f	182	17	1.1	4841	2	AF025847	AB023185 Homo sapi
C 110	17	1.1	939	40	AF081062	AF081062 Eulemur m	183	17	1.1	5705	9	HUMTGFAB1	AF025847 Myxococcu
C 111	17	1.1	939	40	AF081063	AF081063 Eulemur m	184	17	1.1	5705	34	ACAMY0IHC	M34057 Human trans
C 112	17	1.1	939	40	AF081064	AF081064 Eulemur m	185	17	1.1	5743	7	AB007875	M30780 A.castellan
C 113	17	1.1	939	40	AF081065	AF081065 Anopheles	186	17	1.1	6085	7	CHCRPTE	AB007875 Homo sapi
C 114	17	1.1	1079	13	CNS011Z4	AL116385 Anopheles	187	17	1.1	6305	35	AF053358	X61598 C.reiflexa C
C 115	17	1.1	1090	3	OY116017	Y116017 Onchocerca	188	17	1.1	6798	9	AF185277	AF185277 Homo sapi
C 116	17	1.1	1123	34	AF063484	AF063484 Rhizobium	189	17	1.1	6794	40	S60080	S60080 DNA polymer
C 117	17	1.1	1189	9	GCIRBPX1	Z11805 G.crassidicu	190	17	1.1	6858	11	HSU49356	U49356 Human DNA p
C 118	17	1.1	1231	8	AF036301	AF036301 Arabidops	191	17	1.1	6912	10	HUMDNPOLCS	L05561 Homo saplen
C 119	17	1.1	1303	9	HUMCMOS	J00119 human humos	192	17	1.1	7055	42	AC013099	AC013099 Drosophill
C 120	17	1.1	1328	1	STMAFR	D49782 Streptomyce	193	17	1.1	7332	9	HUMTANI	M73380 Human TAN-1
C 121	17	1.1	1412	8	AF020561	AF020561 Cupressus	194	17	1.1	9461	1	STMGPLYZ	L37531 Streptomyce

Thu May 11 11:19:41 2000

us-09-080-127-1.oligo.rge

Page 3

195	17	1.1	10065	2	AE000799	AE000799 Methanoba
196	17	1.1	10675	2	AE001043	AE001043 Archaeogl
197	17	1.1	12433	2	ATUCELABCE	1386005 Archaeobact
198	17	1.1	12581	2	SC936	AL049763 Streptomy
199	17	1.1	13370	2	AE001233	AE001233 Treponema
200	17	1.1	13532	2	U32760	U32760 Haemophilus
201	17	1.1	14896	10	HSLDLRL	X13916 human mRNA
202	17	1.1	14896	13	G28611	G28611 human SMS S
203	17	1.1	15213	2	AE001808	AE001808 Thermotog
204	17	1.1	15807	2	MEU72662	MEU72662 Methylobact
205	17	1.1	25275	7	AB019225	AB019225 Arabidops
206	17	1.1	25416	42	AC013100	AC013100 Drosophill
207	17	1.1	25468	34	CER14F7	281503 Caenorhabd1
208	17	1.1	28835	10	HS444G9	298358 Human DNA s
209	17	1.1	30590	1	SC4G2	AL031371 Streptomy
210	17	1.1	30981	7	SCCHRITG	221487 S.cerevisia
211	17	1.1	32421	43	AC017265	AC017265 Drosophill
212	17	1.1	32421	7	SCDPB3	X76053 S.cerevisia
213	17	1.1	33977	34	CERF1G4	281016 Caenorhabd1
214	17	1.1	35437	1	SCF43A	U39997 Caenorhabd1
215	17	1.1	36120	34	CERF56D1	AC004789 Homo sapi
216	17	1.1	36136	11	AC004789	AL035206 Streptomy
217	17	1.1	36368	1	SC9B5	254142 S.pombe chr
218	17	1.1	36833	8	SPAC24H6	U41015 Caenorhabd1
219	17	1.1	38099	34	CERF20B6	AC000038 Homo sapi
220	17	1.1	38861	40	AC000038	AC005565 Homo sapi
221	17	1.1	39441	11	AC005565	AC012323 Homo sapi
222	17	1.1	39770	45	AC012323	282066 Caenorhabd1
223	17	1.1	39873	32	CERW3A8	AL012634 Caenorhabd
224	17	1.1	40053	34	CERY3G8C	U19739 Saccharomyc
225	17	1.1	40197	8	YSCIL8084	AL131644 Streptomy
226	17	1.1	40442	1	SC18	AC014006 Drosophill
227	17	1.1	44032	42	AC014006	AC033620 Mus muscu
228	17	1.1	44411	55	AC033620	AC020529 Drosophill
229	17	1.1	44588	44	AC020529	AC005744 , complet
230	17	1.1	46335	40	AC005744	AC017073 Homo sapi
231	17	1.1	47889	43	AC017073	Continuation (4 of
232	17	1.1	52481	32	CERY3D8.3	AC010709 Drosophill
233	17	1.1	52948	56	AC010709	AC006184 Homo sapi
234	17	1.1	53269	41	AC006184	AC014844 Drosophill
235	17	1.1	54445	42	AC014844	AC017592 Drosophill
236	17	1.1	58191	43	AC017592	AL021646 Mycobacte
237	17	1.1	58280	1	MTV014	AC017812 Drosophill
238	17	1.1	61728	43	AC017812	AC022915 Homo sapi
239	17	1.1	66432	54	AC022915	AC017149 Drosophill
240	17	1.1	67657	43	AC017149	AC022776 Mus muscu
241	17	1.1	72209	54	AC022776	AC022983 Homo sapi
242	17	1.1	73643	45	AC022983	AC021936 Homo sapi
243	17	1.1	75183	45	AC021936	AC008568 Homo sapi
244	17	1.1	75767	45	AC008568	AT124045 Sorghum b
245	17	1.1	78195	8	AF124045	AB019236 Arabidops
246	17	1.1	80511	56	AC010573	AE217759 Homo sapi
247	17	1.1	81494	7	AB019236	AC003923 Drosophill
248	17	1.1	82369	44	AF217759	AF166999 Homo sapi
249	17	1.1	82393	34	AC003923	AC019808 Drosophill
250	17	1.1	82532	41	AF166999	ATJ11017 Mus muscu
251	17	1.1	85221	44	AC019808	AB011485 Arabidops
252	17	1.1	86380	12	MMU131017	AC005869 *** SEQE
253	17	1.1	87210	7	AB011485	AC012430 Homo sapi
254	17	1.1	91051	46	AC005869	AC004596 Homo sapi
255	17	1.1	92459	42	AC012430	AC004136 Arabidops
256	17	1.1	92638	40	AC004596	AL022322 Human DNA
257	17	1.1	92744	50	AC004136	AC010114 Drosophill
258	17	1.1	95375	11	HS228A9	AC010350 Homo sapi
259	17	1.1	95992	56	AC010114	AC002020 Drosophill
260	17	1.1	97060	44	AC010350	AC005267 Drosophill
261	17	1.1	98576	35	AC020220	AF069442 Arabidops
262	17	1.1	98645	44	AC005267	AC013845 Drosophill
263	17	1.1	98913	8	AF069442	AC014803 Drosophill
264	17	1.1	99699	42	AC013845	AC014357 Drosophill
265	17	1.1	103054	42	AC014803	AL139144 Homo sapi
266	17	1.1	103746	42	AC014357	
267	17	1.1	104387	33	AL139144	
268	17	1.1	105733	8	AC012561	AC012561 Arabidops
269	17	1.1	106209	1	D90903	D90903 Synchocyst
270	17	1.1	106696	40	HUAC003661	AC003661 Homo sapi
271	17	1.1	108709	41	AF161516	AF161516 Homo sapi
272	17	1.1	109522	41	AF161516	Continuation (3 of
273	17	1.1	110000	32	CERY3D8.2	294158 Caenorhabd1
274	17	1.1	110000	32	CERY3E4.0	Continuation (3 of
275	17	1.1	110000	32	CERY3G6.2	Continuation (3 of
276	17	1.1	110000	32	HS5171M.2	Continuation (3 of
277	17	1.1	110000	37	HSJ342K12_2	Continuation (3 of
278	17	1.1	110000	57	AC010187	AC009251 Drosophill
279	17	1.1	111328	54	AC009251	AC009251 Drosophill
280	17	1.1	112228	42	AC015656	AC015656 Homo sapi
281	17	1.1	112597	10	HS82L10	AL078614 Homo sapi
282	17	1.1	112797	33	AL158727	AL158727 Homo sapi
283	17	1.1	113253	42	AC010121	AC0021711 Arabidops
284	17	1.1	114738	8	AC002130	AC008632 Homo sapi
285	17	1.1	114860	44	HS1800J21	AL109955 Homo sapi
286	17	1.1	115294	33	AC020574	AC020574 Homo sapi
287	17	1.1	115294	44	AC020574	AC012676 Homo sapi
288	17	1.1	116119	54	AC012676	AC013527 Homo sapi
289	17	1.1	116515	43	AC013527	AL034452 Human DNA
290	17	1.1	118431	10	HS682J15	AL021711 Arabidops
291	17	1.1	119111	8	ATF13C5	AC005588 Homo sapi
292	17	1.1	119483	40	AC005588	AC006463 Homo sapi
293	17	1.1	120984	40	AC006463	AC010338 Homo sapi
294	17	1.1	121823	43	AC010338	AC008360 Drosophill
295	17	1.1	121857	41	AC008360	AC008226 Drosophill
296	17	1.1	122405	55	AC008226	AC006343 Homo sapi
297	17	1.1	124876	40	AC006343	AC011099 Homo sapi
298	17	1.1	126101	45	AC011099	AC014963 Drosophill
299	17	1.1	126121	42	AC014963	DMBRI8C8
300	17	1.1	126340	32	DMBRI8C8	AC005986 Drosophill
301	17	1.1	128465	46	AC005986	AC020029 Drosophill
302	17	1.1	128465	44	AC020029	AC006422 Drosophill
303	17	1.1	129605	35	AC006422	AC008422 Drosophill
304	17	1.1	130212	43	AC008422	AC013763 Homo sapi
305	17	1.1	130385	32	AC013763	HSAC000383
306	17	1.1	133469	42	HSAC000383	AC004542 Homo sapi
307	17	1.1	134914	51	AC004542	AL138889 Homo sapi
308	17	1.1	139385	33	AL138889	AC010048 Drosophill
309	17	1.1	140283	56	AC010048	AC021022 Homo sapi
310	17	1.1	141248	58	AC021022	AL137218 Homo sapi
311	17	1.1	141503	33	AL137218	AC007743 Homo sapi
312	17	1.1	141836	40	AC007743	AC013547 Homo sapi
313	17	1.1	144313	55	AC013547	AF001548 Human Chr
314	17	1.1	144370	51	HUAF001548	AC003668 Homo sapi
315	17	1.1	145831	40	AC003668	AC018943 Homo sapi
316	17	1.1	146058	45	AC018943	AL117329 Human DNA
317	17	1.1	146325	41	HSBA191L9	AC016404 Homo sapi
318	17	1.1	146438	42	AC011167	AL139423 Homo sapi
319	17	1.1	147025	54	AC016404	AC013533 Homo sapi
320	17	1.1	148381	33	AL139423	AC016358 Homo sapi
321	17	1.1	149932	53	AL139423	AC022799 Homo sapi
322	17	1.1	152589	44	AC016358	AC007859 Homo sapi
323	17	1.1	152769	44	AC022799	AC011199 Homo sapi
324	17	1.1	152910	55	AC011199	AC01234 Homo sapi
325	17	1.1	152925	41	AC007859	AF042090 Homo sapi
326	17	1.1	153665	42	AC015671	AC023478 Homo sapi
327	17	1.1	154754	11	AF042090	AC022258 Homo sapi
328	17	1.1	155407	11	AF042090	AL035070 Human DNA
329	17	1.1	155432	55	AC023478	AC010162 Homo sapi
330	17	1.1	155789	52	AC023478	AC012131 Homo sapi
331	17	1.1	156383	52	AC022258	AC018574 Homo sapi
332	17	1.1	156646	40	AC015671	AF081810 Lymantria
333	17	1.1	156666	14	AF081810	AL137782 Homo sapi
334	17	1.1	158149	55	AC010162	AL137782 Homo sapi
335	17	1.1	158660	42	AC012131	AC018382 Homo sapi
336	17	1.1	160513	44	AC018574	
337	17	1.1	160513	44	AC018574	
338	17	1.1	161046	16	AF081810	
339	17	1.1	161396	33	AL137782	
340	17	1.1	161879	44	AC018382	

C 341	17	1.1	162405	40	AC002352	AC002352 Homo sapi	414	17	1.1	212752	33	AL137066	AL137066 Homo sapi
C 342	17	1.1	162700	43	AC013759	AC013759 Homo sapi	415	17	1.1	213732	2	AE001862	AE001862 Deinococc
C 343	17	1.1	162707	33	AL138776	AL138776 Homo sapi	C 416	17	1.1	215416	57	AC008081	AC008081 Homo sapi
C 344	17	1.1	162912	11	HSAC002067	AC002067 Human BAC	C 417	17	1.1	215719	42	AC009825	AC009825 Homo sapi
C 345	17	1.1	163194	42	AC011985	AC011985 Homo sapi	C 418	17	1.1	215719	42	AC009825	AC009825 Homo sapi
C 346	17	1.1	164952	43	AC013743	AC013743 Homo sapi	C 419	17	1.1	217644	57	AC012036	AC012036 Homo sapi
C 347	17	1.1	164960	35	AC021971	AC021971 Homo sapi	C 420	17	1.1	220936	44	AC020256	AC020256 Drosophill
C 348	17	1.1	164993	33	AL138737	AL138737 Homo sapi	C 421	17	1.1	224416	54	AC015933	AC015933 Homo sapi
C 349	17	1.1	165010	11	HSDB839B4	AL109754 Human DNA	C 422	17	1.1	224568	57	AC007619	AC007619 Homo sapi
C 350	17	1.1	166200	44	AC004816	AC004816 Homo sapi	C 423	17	1.1	229349	33	AC010287	AC010287 Homo sapi
C 351	17	1.1	166330	11	AC011505	AC011505 Homo sapi	C 424	17	1.1	234328	43	AL136439	AL136439 Homo sapi
C 352	17	1.1	166701	11	AC005747	AC005747 Homo sapi	C 425	17	1.1	243064	52	AC007083	AC007083 Drosophill
C 353	17	1.1	167625	43	AC015959	AC015959 Homo sapi	C 426	17	1.1	247275	32	AL135939	AL135939 Homo sapi
C 354	17	1.1	168626	32	AP001034	AP001034 Homo sapi	C 427	17	1.1	247695	1	AP000064	AP000064 Aetopyrum
C 355	17	1.1	168679	54	AC021029	AC021029 Homo sapi	C 428	17	1.1	254877	41	AC006904	AC006904 Caenorhab
C 356	17	1.1	168833	43	AC015565	AC015565 Homo sapi	C 429	17	1.1	257515	32	CEV51H1	292821 Caenorhab
C 357	17	1.1	168833	43	AC015565	AC015565 Homo sapi	C 430	17	1.1	258984	35	AE001274	AE001274 Leishmani
C 358	17	1.1	168986	40	AC008013	AC008013 Homo sapi	C 431	17	1.1	274508	41	AC010530	AC010530 Homo sapi
C 359	17	1.1	169424	42	AC013934	AC013934 Homo sapi	C 432	17	1.1	292390	44	AC006900	AC006900 Caenorhab
C 360	17	1.1	169516	42	AC013934	AC013934 Homo sapi	C 433	17	1.1	299202	41	AC006792	AC006792 Caenorhab
C 361	17	1.1	169653	45	AC021892	AC021892 Oryza sat	C 434	16	1.1	312267	41	AC006785	AC006785 Caenorhab
C 362	17	1.1	169850	45	AC016733	AC016733 Homo sapi	C 435	16	1.1	147	9	AB021476	AB021476 Homo sapi
C 363	17	1.1	170118	51	AC012384	AC012384 Homo sapi	C 436	16	1.1	147	9	AB021477	AB021477 Homo sapi
C 364	17	1.1	171501	32	AP001105	AP001105 Homo sapi	C 437	16	1.1	162	11	AF066102	AF066102 Homo sapi
C 365	17	1.1	171510	32	AL136321	AL136321 Homo sapi	C 438	16	1.1	162	11	AF066480	AF066480 Homo sapi
C 366	17	1.1	172470	56	AC023167	AC023167 Mus muscu	C 439	16	1.1	162	11	AF066659	AF066659 Homo sapi
C 367	17	1.1	172675	41	AC007635	AC007635 Arabidops	C 440	16	1.1	163	11	AF066141	AF066141 Homo sapi
C 368	17	1.1	172943	42	AC013293	AC013293 Homo sapi	C 441	16	1.1	163	11	AF066151	AF066151 Homo sapi
C 369	17	1.1	174079	43	AC016903	AC016903 Homo sapi	C 442	16	1.1	167	11	AF066325	AF066325 Homo sapi
C 370	17	1.1	175504	40	AC010183	AC010183 Homo sapi	C 443	16	1.1	167	11	AF066630	AF066630 Homo sapi
C 371	17	1.1	175775	40	AC006947	AC006947 Homo sapi	C 444	16	1.1	167	11	AF066425	AF066425 Homo sapi
C 372	17	1.1	175775	40	AC006947	AC006947 Homo sapi	C 445	16	1.1	167	11	AF066465	AF066465 Homo sapi
C 373	17	1.1	176186	45	AC006925	AC006925 Homo sapi	C 446	16	1.1	167	11	AF066489	AF066489 Homo sapi
C 374	17	1.1	176838	45	AC020675	AC020675 Homo sapi	C 447	16	1.1	167	11	AF066501	AF066501 Homo sapi
C 375	17	1.1	176854	32	CNS01DVU	AL136039 Homo sapi	C 448	16	1.1	167	11	AF066526	AF066526 Homo sapi
C 376	17	1.1	176983	43	AC011128	AC011128 Homo sapi	C 449	16	1.1	167	11	AF066701	AF066701 Homo sapi
C 377	17	1.1	177740	35	AC005714	AC005714 Drosophill	C 450	16	1.1	170	10	S62710	S62710 [variable c
C 378	17	1.1	177936	45	AC020667	AC020667 Homo sapi	C 451	16	1.1	172	11	AF066195	AF066195 Homo sapi
C 379	17	1.1	179688	45	AC015756	AC015756 Homo sapi	C 452	16	1.1	172	11	AF066498	AF066498 Homo sapi
C 380	17	1.1	178555	41	AC009858	AC009858 Homo sapi	C 453	16	1.1	172	11	AF066500	AF066500 Homo sapi
C 381	17	1.1	178772	55	AC016180	AC016180 Homo sapi	C 454	16	1.1	172	11	AF066514	AF066514 Homo sapi
C 382	17	1.1	179132	35	AC006414	AC006414 Drosophill	C 455	16	1.1	173	11	AF066644	AF066644 Homo sapi
C 383	17	1.1	179854	40	AC007437	AC007437 Homo sapi	C 456	16	1.1	177	11	AF066599	AF066599 Homo sapi
C 384	17	1.1	179855	42	AC015758	AC015758 Homo sapi	C 457	16	1.1	177	11	AF066683	AF066683 Homo sapi
C 385	17	1.1	179907	42	AC009852	AC009852 Homo sapi	C 458	16	1.1	178	9	HSU08084	HSU08084 Human
C 386	17	1.1	180418	54	AC017095	AC017095 Homo sapi	C 459	16	1.1	179	9	HSU08070	HSU08070 Human
C 387	17	1.1	182076	52	AC010158	AC010158 Homo sapi	C 460	16	1.1	179	11	AF066100	AF066100 Homo sapi
C 388	17	1.1	183059	32	CNS01DRP	AL117692 Homo sapi	C 461	16	1.1	179	11	AF066314	AF066314 Homo sapi
C 389	17	1.1	183249	40	AC004828	AC004828 Homo sapi	C 462	16	1.1	179	11	AF066328	AF066328 Homo sapi
C 390	17	1.1	184204	44	AC011359	AC011359 Homo sapi	C 463	16	1.1	179	11	AF066405	AF066405 Homo sapi
C 391	17	1.1	184886	11	AC005358	AC005358 Homo sapi	C 464	16	1.1	179	11	AF066407	AF066407 Homo sapi
C 392	17	1.1	185363	44	AC016673	AC016673 Homo sapi	C 465	16	1.1	179	11	AF066408	AF066408 Homo sapi
C 393	17	1.1	185759	33	HSDB559A3	AL117348 Homo sapi	C 466	16	1.1	179	11	AF066410	AF066410 Homo sapi
C 394	17	1.1	185959	54	AC021028	AC021028 Homo sapi	C 467	16	1.1	179	11	AF066412	AF066412 Homo sapi
C 395	17	1.1	187306	43	AC013652	AC013652 Homo sapi	C 468	16	1.1	179	11	AF066414	AF066414 Homo sapi
C 396	17	1.1	188354	45	AC018949	AC018949 Homo sapi	C 469	16	1.1	179	11	AF066420	AF066420 Homo sapi
C 397	17	1.1	188915	44	AC016975	AC016975 Homo sapi	C 470	16	1.1	179	11	AF066583	AF066583 Homo sapi
C 398	17	1.1	189039	52	AC017012	AC017012 Homo sapi	C 471	16	1.1	180	9	HSU08054	HSU08054 Human
C 399	17	1.1	189228	44	AC008876	AC008876 Homo sapi	C 472	16	1.1	180	9	HSU08055	HSU08055 Human
C 400	17	1.1	189271	55	AC013254	AC013254 Drosophill	C 473	16	1.1	180	9	HSU08056	HSU08056 Human
C 401	17	1.1	190142	57	AC022246	AC022246 Homo sapi	C 474	16	1.1	180	9	HSU08057	HSU08057 Human
C 402	17	1.1	190362	52	AC016924	AC016924 Homo sapi	C 475	16	1.1	180	9	HSU08058	HSU08058 Human
C 403	17	1.1	192036	34	AC012415	AC012415 Homo sapi	C 476	16	1.1	180	9	HSU08059	HSU08059 Human
C 404	17	1.1	194525	53	AL137121	AL137121 Homo sapi	C 477	16	1.1	180	9	HSU08060	HSU08060 Human
C 405	17	1.1	196098	41	AC011194	AC011194 Mus muscu	C 478	16	1.1	180	9	HSU08062	HSU08062 Human
C 406	17	1.1	196287	10	CNS0000B	AL049829 Human chr	C 479	16	1.1	180	9	HSU08063	HSU08063 Human
C 407	17	1.1	196451	32	CNS01DV0	AL133444 Homo sapi	C 480	16	1.1	180	9	HSU08064	HSU08064 Human
C 408	17	1.1	202370	53	AC019127	AC019127 Homo sapi	C 481	16	1.1	180	9	HSU08065	HSU08065 Human
C 409	17	1.1	203495	45	AC022031	AC022031 Homo sapi	C 482	16	1.1	180	9	HSU08066	HSU08066 Human
C 410	17	1.1	208079	54	AC013391	AC013391 Homo sapi	C 483	16	1.1	180	9	HSU08067	HSU08067 Human
C 411	17	1.1	208628	43	AC021332	AC021332 Homo sapi	C 484	16	1.1	180	9	HSU08068	HSU08068 Human
C 412	17	1.1	210213	53	AC016736	AC016736 Homo sapi	C 485	16	1.1	180	9	HSU08071	HSU08071 Human
C 413	17	1.1	211626	42	AC012250	AC012250 Homo sapi	C 486	16	1.1	180	9	HSU08072	HSU08072 Human
C 413	17	1.1	212420	32	CNS01DRB	AL110504 Homo sapi							

487	16	1.1	180	9	HSU08075	U08075 Human Isola	560	16	1.1	183	11	AF066145	AF066145 Homo sapi
488	16	1.1	180	9	HSU08076	U08076 Human Isola	561	16	1.1	183	11	AF066210	AF066210 Homo sapi
489	16	1.1	180	9	HSU08077	U08077 Human Isola	562	16	1.1	183	11	AF066354	AF066354 Homo sapi
490	16	1.1	180	9	HSU08078	U08078 Human Isola	563	16	1.1	183	11	AF066401	AF066401 Homo sapi
491	16	1.1	180	9	HSU08079	U08079 Human Isola	564	16	1.1	183	11	AF066406	AF066406 Homo sapi
492	16	1.1	180	9	HSU08080	U08080 Human Isola	565	16	1.1	183	11	AF066441	AF066441 Homo sapi
493	16	1.1	180	9	HSU08082	U08082 Human Isola	566	16	1.1	183	11	AF066481	AF066481 Homo sapi
494	16	1.1	180	9	HSU08083	U08083 Human Isola	567	16	1.1	183	11	AF066530	AF066530 Homo sapi
495	16	1.1	180	9	HSU08085	U08085 Human Isola	568	16	1.1	183	11	AF066566	AF066566 Homo sapi
496	16	1.1	180	9	HSU08086	U08086 Human Isola	569	16	1.1	183	11	AF066656	AF066656 Homo sapi
497	16	1.1	180	9	HSU08087	U08087 Human Isola	570	16	1.1	183	11	AF066769	AF066769 Homo sapi
498	16	1.1	180	9	HSU08088	U08088 Human Isola	571	16	1.1	183	11	AF066770	AF066770 Homo sapi
499	16	1.1	180	9	HSU08089	U08089 Human Isola	572	16	1.1	183	11	AF066772	AF066772 Homo sapi
500	16	1.1	180	9	HSU08090	U08090 Human Isola	573	16	1.1	183	11	AF066773	AF066773 Homo sapi
501	16	1.1	181	11	AF066251	AF066251 Homo sapi	574	16	1.1	183	11	AF066783	AF066783 Homo sapi
502	16	1.1	181	11	AF066404	AF066404 Homo sapi	575	16	1.1	183	11	AF066784	AF066784 Homo sapi
503	16	1.1	181	11	AF066468	AF066468 Homo sapi	576	16	1.1	184	11	AF066461	AF066461 Homo sapi
504	16	1.1	181	11	AF066645	AF066645 Homo sapi	577	16	1.1	184	11	AF066476	AF066476 Homo sapi
505	16	1.1	182	9	AB026773	AB026773 Homo sapi	578	16	1.1	184	11	AF066503	AF066503 Homo sapi
506	16	1.1	182	9	AB026774	AB026774 Homo sapi	579	16	1.1	184	11	AF066781	AF066781 Homo sapi
507	16	1.1	182	9	AB026776	AB026776 Homo sapi	580	16	1.1	184	11	AF066786	AF066786 Homo sapi
508	16	1.1	182	9	AB026778	AB026778 Homo sapi	581	16	1.1	184	11	AF066787	AF066787 Homo sapi
509	16	1.1	182	9	AB026779	AB026779 Homo sapi	582	16	1.1	185	9	AB021474	AB021474 Homo sapi
510	16	1.1	182	9	AB026780	AB026780 Homo sapi	583	16	1.1	185	9	AB021482	AB021482 Homo sapi
511	16	1.1	182	9	AB026781	AB026781 Homo sapi	584	16	1.1	185	9	AB021489	AB021489 Homo sapi
512	16	1.1	182	9	AB026782	AB026782 Homo sapi	585	16	1.1	185	11	AF066252	AF066252 Homo sapi
513	16	1.1	182	9	AB026783	AB026783 Homo sapi	586	16	1.1	185	11	AF066417	AF066417 Homo sapi
514	16	1.1	182	9	AB026784	AB026784 Homo sapi	587	16	1.1	185	11	AF066692	AF066692 Homo sapi
515	16	1.1	182	9	AB026786	AB026786 Homo sapi	588	16	1.1	185	11	AF066775	AF066775 Homo sapi
516	16	1.1	182	9	AB026787	AB026787 Homo sapi	589	16	1.1	186	9	AB026777	AB026777 Homo sapi
517	16	1.1	182	9	AB026788	AB026788 Homo sapi	590	16	1.1	186	9	AB026785	AB026785 Homo sapi
518	16	1.1	182	9	AB026789	AB026789 Homo sapi	591	16	1.1	186	11	AF066113	AF066113 Homo sapi
519	16	1.1	182	9	AB026790	AB026790 Homo sapi	592	16	1.1	186	11	AF066115	AF066115 Homo sapi
520	16	1.1	182	9	AB026791	AB026791 Homo sapi	593	16	1.1	186	11	AF066118	AF066118 Homo sapi
521	16	1.1	182	9	AB026792	AB026792 Homo sapi	594	16	1.1	186	11	AF066167	AF066167 Homo sapi
522	16	1.1	182	9	AB026793	AB026793 Homo sapi	595	16	1.1	186	11	AF066185	AF066185 Homo sapi
523	16	1.1	182	9	AB026794	AB026794 Homo sapi	596	16	1.1	186	11	AF066188	AF066188 Homo sapi
524	16	1.1	182	9	AB026795	AB026795 Homo sapi	597	16	1.1	186	11	AF066256	AF066256 Homo sapi
525	16	1.1	182	9	AB026797	AB026797 Homo sapi	598	16	1.1	186	11	AF066259	AF066259 Homo sapi
526	16	1.1	182	9	AB026798	AB026798 Homo sapi	599	16	1.1	186	11	AF066261	AF066261 Homo sapi
527	16	1.1	182	9	AB026800	AB026800 Homo sapi	600	16	1.1	186	11	AF066275	AF066275 Homo sapi
528	16	1.1	182	9	AB026801	AB026801 Homo sapi	601	16	1.1	186	11	AF066331	AF066331 Homo sapi
529	16	1.1	182	11	AF066090	AF066090 Homo sapi	602	16	1.1	186	11	AF066373	AF066373 Homo sapi
530	16	1.1	182	11	AF066097	AF066097 Homo sapi	603	16	1.1	186	11	AF066388	AF066388 Homo sapi
531	16	1.1	182	11	AF066108	AF066108 Homo sapi	604	16	1.1	186	11	AF066455	AF066455 Homo sapi
532	16	1.1	182	11	AF066123	AF066123 Homo sapi	605	16	1.1	186	11	AF066482	AF066482 Homo sapi
533	16	1.1	182	11	AF066138	AF066138 Homo sapi	606	16	1.1	186	11	AF066484	AF066484 Homo sapi
534	16	1.1	182	11	AF066166	AF066166 Homo sapi	607	16	1.1	186	11	AF066486	AF066486 Homo sapi
535	16	1.1	182	11	AF066186	AF066186 Homo sapi	608	16	1.1	186	11	AF066492	AF066492 Homo sapi
536	16	1.1	182	11	AF066190	AF066190 Homo sapi	609	16	1.1	186	11	AF066494	AF066494 Homo sapi
537	16	1.1	182	11	AF066246	AF066246 Homo sapi	610	16	1.1	186	11	AF066496	AF066496 Homo sapi
538	16	1.1	182	11	AF066248	AF066248 Homo sapi	611	16	1.1	186	11	AF066505	AF066505 Homo sapi
539	16	1.1	182	11	AF066342	AF066342 Homo sapi	612	16	1.1	186	11	AF066515	AF066515 Homo sapi
540	16	1.1	182	11	AF066444	AF066444 Homo sapi	613	16	1.1	186	11	AF066516	AF066516 Homo sapi
541	16	1.1	182	11	AF066445	AF066445 Homo sapi	614	16	1.1	186	11	AF066519	AF066519 Homo sapi
542	16	1.1	182	11	AF066454	AF066454 Homo sapi	615	16	1.1	186	11	AF066520	AF066520 Homo sapi
543	16	1.1	182	11	AF066457	AF066457 Homo sapi	616	16	1.1	186	11	AF066522	AF066522 Homo sapi
544	16	1.1	182	11	AF066463	AF066463 Homo sapi	617	16	1.1	186	11	AF066524	AF066524 Homo sapi
545	16	1.1	182	11	AF066470	AF066470 Homo sapi	618	16	1.1	186	11	AF066525	AF066525 Homo sapi
546	16	1.1	182	11	AF066485	AF066485 Homo sapi	619	16	1.1	186	11	AF066535	AF066535 Homo sapi
547	16	1.1	182	11	AF066487	AF066487 Homo sapi	620	16	1.1	186	11	AF066536	AF066536 Homo sapi
548	16	1.1	182	11	AF066491	AF066491 Homo sapi	621	16	1.1	186	11	AF066531	AF066531 Homo sapi
549	16	1.1	182	11	AF066510	AF066510 Homo sapi	622	16	1.1	186	11	AF066606	AF066606 Homo sapi
550	16	1.1	182	11	AF066511	AF066511 Homo sapi	623	16	1.1	186	11	AF066635	AF066635 Homo sapi
551	16	1.1	182	11	AF066518	AF066518 Homo sapi	624	16	1.1	186	11	AF066637	AF066637 Homo sapi
552	16	1.1	182	11	AF066552	AF066552 Homo sapi	625	16	1.1	187	11	AF066692	AF066692 Homo sapi
553	16	1.1	182	11	AF066553	AF066553 Homo sapi	626	16	1.1	187	11	AF066694	AF066694 Homo sapi
554	16	1.1	182	11	AF066574	AF066574 Homo sapi	627	16	1.1	187	11	AF066695	AF066695 Homo sapi
555	16	1.1	182	11	AF066591	AF066591 Homo sapi	628	16	1.1	187	11	AF0666103	AF0666103 Homo sapi
556	16	1.1	182	11	AF066602	AF066602 Homo sapi	629	16	1.1	187	11	AF0666104	AF0666104 Homo sapi
557	16	1.1	182	11	AF066633	AF066633 Homo sapi	630	16	1.1	187	11	AF0666110	AF0666110 Homo sapi
558	16	1.1	182	11	AF066723	AF066723 Homo sapi	631	16	1.1	187	11	AF0666132	AF0666132 Homo sapi
559	16	1.1	183	11	AF066116	AF066116 Homo sapi	632	16	1.1	187	11	AF0666153	AF0666153 Homo sapi

633	16	1.1	187	11	AF066150	AF066160	Homo	sap1	706	16	1.1	187	11	HS047145	U47145	Human	mtoc
634	16	1.1	187	11	AF066171	AF066171	Homo	sap1	707	16	1.1	187	11	HS047146	U47146	Human	mtoc
635	16	1.1	187	11	AF066176	AF066176	Homo	sap1	708	16	1.1	187	11	HS047147	U47147	Human	mtoc
636	16	1.1	187	11	AF066184	AF066184	Homo	sap1	709	16	1.1	187	11	HS047148	U47148	Human	mtoc
637	16	1.1	187	11	AF066187	AF066187	Homo	sap1	710	16	1.1	187	11	HS047149	U47149	Human	mtoc
638	16	1.1	187	11	AF066193	AF066193	Homo	sap1	711	16	1.1	187	11	HS047150	U47150	Human	mtoc
639	16	1.1	187	11	AF066208	AF066208	Homo	sap1	712	16	1.1	187	11	HS047151	U47151	Human	mtoc
640	16	1.1	187	11	AF066211	AF066211	Homo	sap1	713	16	1.1	187	11	HS047152	U47152	Human	mtoc
641	16	1.1	187	11	AF066212	AF066212	Homo	sap1	714	16	1.1	187	11	HS047153	U47153	Human	mtoc
642	16	1.1	187	11	AF066215	AF066215	Homo	sap1	715	16	1.1	187	11	HS047154	U47154	Human	mtoc
643	16	1.1	187	11	AF066241	AF066241	Homo	sap1	716	16	1.1	187	11	HS047154	U47154	Human	mtoc
644	16	1.1	187	11	AF066249	AF066249	Homo	sap1	717	16	1.1	187	11	HS047193	U47193	Human	mtoc
645	16	1.1	187	11	AF066311	AF066311	Homo	sap1	718	16	1.1	187	11	HS047207	U47207	Human	mtoc
646	16	1.1	187	11	AF066322	AF066322	Homo	sap1	719	16	1.1	187	11	HS047208	U47208	Human	mtoc
647	16	1.1	187	11	AF066327	AF066327	Homo	sap1	720	16	1.1	187	11	HS047210	U47210	Human	mtoc
648	16	1.1	187	11	AF066335	AF066335	Homo	sap1	721	16	1.1	187	11	HS047211	U47211	Human	mtoc
649	16	1.1	187	11	AF066402	AF066402	Homo	sap1	722	16	1.1	187	11	HS047212	U47212	Human	mtoc
650	16	1.1	187	11	AF066409	AF066409	Homo	sap1	723	16	1.1	187	11	HS047236	U47236	Human	mtoc
651	16	1.1	187	11	AF066413	AF066413	Homo	sap1	724	16	1.1	187	11	HS047237	U47237	Human	mtoc
652	16	1.1	187	11	AF066415	AF066415	Homo	sap1	725	16	1.1	187	11	HS047238	U47238	Human	mtoc
653	16	1.1	187	11	AF066419	AF066419	Homo	sap1	726	16	1.1	187	11	HS047239	U47239	Human	mtoc
654	16	1.1	187	11	AF066421	AF066421	Homo	sap1	727	16	1.1	187	11	HS047240	U47240	Human	mtoc
655	16	1.1	187	11	AF066422	AF066422	Homo	sap1	728	16	1.1	187	11	HS047235	U47235	Human	mtoc
656	16	1.1	187	11	AF066423	AF066423	Homo	sap1	729	16	1.1	187	11	HS047266	U47266	Human	mtoc
657	16	1.1	187	11	AF066434	AF066434	Homo	sap1	730	16	1.1	187	11	HS047270	U47270	Human	mtoc
658	16	1.1	187	11	AF066434	AF066434	Homo	sap1	731	16	1.1	188	11	AF066306	AF066306	Human	mtoc
659	16	1.1	187	11	AF066434	AF066434	Homo	sap1	732	16	1.1	188	11	AF066567	AF066567	Homo	sap1
660	16	1.1	187	11	AF066449	AF066449	Homo	sap1	733	16	1.1	189	11	AB026775	AB026775	Homo	sap1
661	16	1.1	187	11	AF066483	AF066483	Homo	sap1	734	16	1.1	189	11	AF066336	AF066336	Homo	sap1
662	16	1.1	187	11	AF066490	AF066490	Homo	sap1	735	16	1.1	189	11	AF066338	AF066338	Homo	sap1
663	16	1.1	187	11	AF066513	AF066513	Homo	sap1	736	16	1.1	189	11	AF066366	AF066366	Homo	sap1
664	16	1.1	187	11	AF066527	AF066527	Homo	sap1	737	16	1.1	189	11	AF066403	AF066403	Homo	sap1
665	16	1.1	187	11	AF066530	AF066530	Homo	sap1	738	16	1.1	189	11	AF066474	AF066474	Homo	sap1
666	16	1.1	187	11	AF066534	AF066534	Homo	sap1	739	16	1.1	189	11	AF066523	AF066523	Homo	sap1
667	16	1.1	187	11	AF066538	AF066538	Homo	sap1	740	16	1.1	189	11	AF066561	AF066561	Homo	sap1
668	16	1.1	187	11	AF066542	AF066542	Homo	sap1	741	16	1.1	189	11	AF066563	AF066563	Homo	sap1
669	16	1.1	187	11	AF066546	AF066546	Homo	sap1	742	16	1.1	189	11	AF066675	AF066675	Homo	sap1
670	16	1.1	187	11	AF066553	AF066553	Homo	sap1	743	16	1.1	189	11	AF066679	AF066679	Homo	sap1
671	16	1.1	187	11	AF066557	AF066557	Homo	sap1	744	16	1.1	190	11	AF066698	AF066698	Homo	sap1
672	16	1.1	187	11	AF066565	AF066565	Homo	sap1	745	16	1.1	190	11	AF066117	AF066117	Homo	sap1
673	16	1.1	187	11	AF066566	AF066566	Homo	sap1	746	16	1.1	190	11	AF066126	AF066126	Homo	sap1
674	16	1.1	187	11	AF066577	AF066577	Homo	sap1	747	16	1.1	190	11	AF066131	AF066131	Homo	sap1
675	16	1.1	187	11	AF066593	AF066593	Homo	sap1	748	16	1.1	190	11	AF066133	AF066133	Homo	sap1
676	16	1.1	187	11	AF066595	AF066595	Homo	sap1	749	16	1.1	190	11	AF066140	AF066140	Homo	sap1
677	16	1.1	187	11	AF066596	AF066596	Homo	sap1	750	16	1.1	190	11	AF066143	AF066143	Homo	sap1
678	16	1.1	187	11	AF066597	AF066597	Homo	sap1	751	16	1.1	190	11	AF066148	AF066148	Homo	sap1
679	16	1.1	187	11	AF066611	AF066611	Homo	sap1	752	16	1.1	190	11	AF066156	AF066156	Homo	sap1
680	16	1.1	187	11	AF066613	AF066613	Homo	sap1	753	16	1.1	190	11	AF066158	AF066158	Homo	sap1
681	16	1.1	187	11	AF066616	AF066616	Homo	sap1	754	16	1.1	190	11	AF066179	AF066179	Homo	sap1
682	16	1.1	187	11	AF066629	AF066629	Homo	sap1	755	16	1.1	190	11	AF066180	AF066180	Homo	sap1
683	16	1.1	187	11	AF066639	AF066639	Homo	sap1	756	16	1.1	190	11	AF066182	AF066182	Homo	sap1
684	16	1.1	187	11	AF066642	AF066642	Homo	sap1	757	16	1.1	190	11	AF066201	AF066201	Homo	sap1
685	16	1.1	187	11	AF066643	AF066643	Homo	sap1	758	16	1.1	190	11	AF066201	AF066201	Homo	sap1
686	16	1.1	187	11	AF066646	AF066646	Homo	sap1	759	16	1.1	190	11	AF066202	AF066202	Homo	sap1
687	16	1.1	187	11	AF066654	AF066654	Homo	sap1	760	16	1.1	190	11	AF066204	AF066204	Homo	sap1
688	16	1.1	187	11	AF066657	AF066657	Homo	sap1	761	16	1.1	190	11	AF066205	AF066205	Homo	sap1
689	16	1.1	187	11	AF066674	AF066674	Homo	sap1	762	16	1.1	190	11	AF066217	AF066217	Homo	sap1
690	16	1.1	187	11	AF066678	AF066678	Homo	sap1	763	16	1.1	190	11	AF066218	AF066218	Homo	sap1
691	16	1.1	187	11	AF066680	AF066680	Homo	sap1	764	16	1.1	190	11	AF066220	AF066220	Homo	sap1
692	16	1.1	187	11	AF066681	AF066681	Homo	sap1	765	16	1.1	190	11	AF066221	AF066221	Homo	sap1
693	16	1.1	187	11	AF066685	AF066685	Homo	sap1	766	16	1.1	190	11	AF066223	AF066223	Homo	sap1
694	16	1.1	187	11	AF066687	AF066687	Homo	sap1	767	16	1.1	190	11	AF066225	AF066225	Homo	sap1
695	16	1.1	187	11	AF066689	AF066689	Homo	sap1	768	16	1.1	190	11	AF066231	AF066231	Homo	sap1
696	16	1.1	187	11	AF066690	AF066690	Homo	sap1	769	16	1.1	190	11	AF066234	AF066234	Homo	sap1
697	16	1.1	187	11	AF066693	AF066693	Homo	sap1	770	16	1.1	190	11	AF066240	AF066240	Homo	sap1
698	16	1.1	187	11	AF066694	AF066694	Homo	sap1	771	16	1.1	190	11	AF066247	AF066247	Homo	sap1
699	16	1.1	187	11	AF066697	AF066697	Homo	sap1	772	16	1.1	190	11	AF066253	AF066253	Homo	sap1
700	16	1.1	187	11	AF066698	AF066698	Homo	sap1	773	16	1.1	190	11	AF066258	AF066258	Homo	sap1
701	16	1.1	187	11	AF066702	AF066702	Homo	sap1	774	16	1.1	190	11	AF066260	AF066260	Homo	sap1
702	16	1.1	187	11	AF066703	AF066703	Homo	sap1	775	16	1.1	190	11	AF066263	AF066263	Homo	sap1
703	16	1.1	187	11	AF066747	AF066747	Homo	sap1	776	16	1.1	190	11	AF066264	AF066264	Homo	sap1
704	16	1.1	187	11	AF066766	AF066766	Homo	sap1	777	16	1.1	190	11	AF066268	AF066268	Homo	sap1
705	16	1.1	187	11	AF066768	AF066768	Homo	sap1	778	16	1.1	190	11	AF066270	AF066270	Homo	sap1

779	16	1.1	190	11	AF066272	Homo	sapi	852	16	1.1	190	11	AF066528	Homo	sapi	AF066528
780	16	1.1	190	11	AF066273	Homo	sapi	853	16	1.1	190	11	AF066532	Homo	sapi	AF066532
781	16	1.1	190	11	AF066276	Homo	sapi	854	16	1.1	190	11	AF066533	Homo	sapi	AF066533
782	16	1.1	190	11	AF066277	Homo	sapi	855	16	1.1	190	11	AF066537	Homo	sapi	AF066537
783	16	1.1	190	11	AF066282	Homo	sapi	856	16	1.1	190	11	AF066544	Homo	sapi	AF066544
784	16	1.1	190	11	AF066288	Homo	sapi	857	16	1.1	190	11	AF066545	Homo	sapi	AF066545
785	16	1.1	190	11	AF066290	Homo	sapi	858	16	1.1	190	11	AF066549	Homo	sapi	AF066549
786	16	1.1	190	11	AF066291	Homo	sapi	859	16	1.1	190	11	AF066554	Homo	sapi	AF066554
787	16	1.1	190	11	AF066302	Homo	sapi	860	16	1.1	190	11	AF066555	Homo	sapi	AF066555
788	16	1.1	190	11	AF066303	Homo	sapi	861	16	1.1	190	11	AF066558	Homo	sapi	AF066558
789	16	1.1	190	11	AF066310	Homo	sapi	862	16	1.1	190	11	AF066559	Homo	sapi	AF066559
790	16	1.1	190	11	AF066313	Homo	sapi	863	16	1.1	190	11	AF066560	Homo	sapi	AF066560
791	16	1.1	190	11	AF066315	Homo	sapi	864	16	1.1	190	11	AF066571	Homo	sapi	AF066571
792	16	1.1	190	11	AF066316	Homo	sapi	865	16	1.1	190	11	AF066572	Homo	sapi	AF066572
793	16	1.1	190	11	AF066317	Homo	sapi	866	16	1.1	190	11	AF066573	Homo	sapi	AF066573
794	16	1.1	190	11	AF066319	Homo	sapi	867	16	1.1	190	11	AF066575	Homo	sapi	AF066575
795	16	1.1	190	11	AF066320	Homo	sapi	868	16	1.1	190	11	AF066576	Homo	sapi	AF066576
796	16	1.1	190	11	AF066324	Homo	sapi	869	16	1.1	190	11	AF066579	Homo	sapi	AF066579
797	16	1.1	190	11	AF066334	Homo	sapi	870	16	1.1	190	11	AF066580	Homo	sapi	AF066580
798	16	1.1	190	11	AF066338	Homo	sapi	871	16	1.1	190	11	AF066585	Homo	sapi	AF066585
799	16	1.1	190	11	AF066339	Homo	sapi	872	16	1.1	190	11	AF066587	Homo	sapi	AF066587
800	16	1.1	190	11	AF066341	Homo	sapi	873	16	1.1	190	11	AF066588	Homo	sapi	AF066588
801	16	1.1	190	11	AF066343	Homo	sapi	874	16	1.1	190	11	AF066589	Homo	sapi	AF066589
802	16	1.1	190	11	AF066347	Homo	sapi	875	16	1.1	190	11	AF066594	Homo	sapi	AF066594
803	16	1.1	190	11	AF066350	Homo	sapi	876	16	1.1	190	11	AF066600	Homo	sapi	AF066600
804	16	1.1	190	11	AF066353	Homo	sapi	877	16	1.1	190	11	AF066601	Homo	sapi	AF066601
805	16	1.1	190	11	AF066355	Homo	sapi	878	16	1.1	190	11	AF066603	Homo	sapi	AF066603
806	16	1.1	190	11	AF066357	Homo	sapi	879	16	1.1	190	11	AF066604	Homo	sapi	AF066604
807	16	1.1	190	11	AF066359	Homo	sapi	880	16	1.1	190	11	AF066607	Homo	sapi	AF066607
808	16	1.1	190	11	AF066360	Homo	sapi	881	16	1.1	190	11	AF066608	Homo	sapi	AF066608
809	16	1.1	190	11	AF066365	Homo	sapi	882	16	1.1	190	11	AF066609	Homo	sapi	AF066609
810	16	1.1	190	11	AF066367	Homo	sapi	883	16	1.1	190	11	AF066610	Homo	sapi	AF066610
811	16	1.1	190	11	AF066370	Homo	sapi	884	16	1.1	190	11	AF066612	Homo	sapi	AF066612
812	16	1.1	190	11	AF066371	Homo	sapi	885	16	1.1	190	11	AF066614	Homo	sapi	AF066614
813	16	1.1	190	11	AF066374	Homo	sapi	886	16	1.1	190	11	AF066615	Homo	sapi	AF066615
814	16	1.1	190	11	AF066375	Homo	sapi	887	16	1.1	190	11	AF066617	Homo	sapi	AF066617
815	16	1.1	190	11	AF066376	Homo	sapi	888	16	1.1	190	11	AF066618	Homo	sapi	AF066618
816	16	1.1	190	11	AF066377	Homo	sapi	889	16	1.1	190	11	AF066620	Homo	sapi	AF066620
817	16	1.1	190	11	AF066378	Homo	sapi	890	16	1.1	190	11	AF066622	Homo	sapi	AF066622
818	16	1.1	190	11	AF066379	Homo	sapi	891	16	1.1	190	11	AF066624	Homo	sapi	AF066624
819	16	1.1	190	11	AF066383	Homo	sapi	892	16	1.1	190	11	AF066625	Homo	sapi	AF066625
820	16	1.1	190	11	AF066385	Homo	sapi	893	16	1.1	190	11	AF066626	Homo	sapi	AF066626
821	16	1.1	190	11	AF066387	Homo	sapi	894	16	1.1	190	11	AF066627	Homo	sapi	AF066627
822	16	1.1	190	11	AF066389	Homo	sapi	895	16	1.1	190	11	AF066628	Homo	sapi	AF066628
823	16	1.1	190	11	AF066390	Homo	sapi	896	16	1.1	190	11	AF066632	Homo	sapi	AF066632
824	16	1.1	190	11	AF066400	Homo	sapi	897	16	1.1	190	11	AF066634	Homo	sapi	AF066634
825	16	1.1	190	11	AF066411	Homo	sapi	898	16	1.1	190	11	AF066636	Homo	sapi	AF066636
826	16	1.1	190	11	AF066416	Homo	sapi	899	16	1.1	190	11	AF066640	Homo	sapi	AF066640
827	16	1.1	190	11	AF066426	Homo	sapi	900	16	1.1	190	11	AF066641	Homo	sapi	AF066641
828	16	1.1	190	11	AF066439	Homo	sapi	901	16	1.1	190	11	AF066647	Homo	sapi	AF066647
829	16	1.1	190	11	AF066448	Homo	sapi	902	16	1.1	190	11	AF066648	Homo	sapi	AF066648
830	16	1.1	190	11	AF066451	Homo	sapi	903	16	1.1	190	11	AF066649	Homo	sapi	AF066649
831	16	1.1	190	11	AF066456	Homo	sapi	904	16	1.1	190	11	AF066650	Homo	sapi	AF066650
832	16	1.1	190	11	AF066458	Homo	sapi	905	16	1.1	190	11	AF066651	Homo	sapi	AF066651
833	16	1.1	190	11	AF066459	Homo	sapi	906	16	1.1	190	11	AF066652	Homo	sapi	AF066652
834	16	1.1	190	11	AF066460	Homo	sapi	907	16	1.1	190	11	AF066655	Homo	sapi	AF066655
835	16	1.1	190	11	AF066462	Homo	sapi	908	16	1.1	190	11	AF066657	Homo	sapi	AF066657
836	16	1.1	190	11	AF066464	Homo	sapi	909	16	1.1	190	11	AF066660	Homo	sapi	AF066660
837	16	1.1	190	11	AF066466	Homo	sapi	910	16	1.1	190	11	AF066661	Homo	sapi	AF066661
838	16	1.1	190	11	AF066467	Homo	sapi	911	16	1.1	190	11	AF066662	Homo	sapi	AF066662
839	16	1.1	190	11	AF066469	Homo	sapi	912	16	1.1	190	11	AF066663	Homo	sapi	AF066663
840	16	1.1	190	11	AF066471	Homo	sapi	913	16	1.1	190	11	AF066664	Homo	sapi	AF066664
841	16	1.1	190	11	AF066477	Homo	sapi	914	16	1.1	190	11	AF066665	Homo	sapi	AF066665
842	16	1.1	190	11	AF066478	Homo	sapi	915	16	1.1	190	11	AF066666	Homo	sapi	AF066666
843	16	1.1	190	11	AF066488	Homo	sapi	916	16	1.1	190	11	AF066667	Homo	sapi	AF066667
844	16	1.1	190	11	AF066493	Homo	sapi	917	16	1.1	190	11	AF066668	Homo	sapi	AF066668
845	16	1.1	190	11	AF066499	Homo	sapi	918	16	1.1	190	11	AF066669	Homo	sapi	AF066669
846	16	1.1	190	11	AF066502	Homo	sapi	919	16	1.1	190	11	AF066670	Homo	sapi	AF066670
847	16	1.1	190	11	AF066506	Homo	sapi	920	16	1.1	190	11	AF066671	Homo	sapi	AF066671
848	16	1.1	190	11	AF066507	Homo	sapi	921	16	1.1	190	11	AF066672	Homo	sapi	AF066672
849	16	1.1	190	11	AF066508	Homo	sapi	922	16	1.1	190	11	AF066676	Homo	sapi	AF066676
850	16	1.1	190	11	AF066512	Homo	sapi	923	16	1.1	190	11	AF066677	Homo	sapi	AF066677
851	16	1.1	190	11	AF066517	Homo	sapi	924	16	1.1	190	11	AF066682	Homo	sapi	AF066682

925	16	1.1	190	11	AF066684	AF066684 Homo sapi
926	16	1.1	190	11	AF066686	AF066686 Homo sapi
927	16	1.1	190	11	AF066688	AF066688 Homo sapi
928	16	1.1	190	11	AF066691	AF066691 Homo sapi
929	16	1.1	190	11	AF066695	AF066695 Homo sapi
930	16	1.1	190	11	AF066696	AF066696 Homo sapi
931	16	1.1	190	11	AF066699	AF066699 Homo sapi
932	16	1.1	190	11	AF066700	AF066700 Homo sapi
933	16	1.1	190	11	AF066704	AF066704 Homo sapi
934	16	1.1	190	11	AF066708	AF066708 Homo sapi
935	16	1.1	190	11	AF066709	AF066709 Homo sapi
936	16	1.1	190	11	AF066710	AF066710 Homo sapi
937	16	1.1	190	11	AF066711	AF066711 Homo sapi
938	16	1.1	190	11	AF066712	AF066712 Homo sapi
939	16	1.1	190	11	AF066713	AF066713 Homo sapi
940	16	1.1	190	11	AF066715	AF066715 Homo sapi
941	16	1.1	190	11	AF066716	AF066716 Homo sapi
942	16	1.1	190	11	AF066717	AF066717 Homo sapi
943	16	1.1	190	11	AF066718	AF066718 Homo sapi
944	16	1.1	190	11	AF066719	AF066719 Homo sapi
945	16	1.1	190	11	AF066720	AF066720 Homo sapi
946	16	1.1	190	11	AF066721	AF066721 Homo sapi
947	16	1.1	190	11	AF066722	AF066722 Homo sapi
948	16	1.1	190	11	AF066724	AF066724 Homo sapi
949	16	1.1	190	11	AF066725	AF066725 Homo sapi
950	16	1.1	190	11	AF066726	AF066726 Homo sapi
951	16	1.1	190	11	AF066727	AF066727 Homo sapi
952	16	1.1	190	11	AF066728	AF066728 Homo sapi
953	16	1.1	190	11	AF066729	AF066729 Homo sapi
954	16	1.1	190	11	AF066730	AF066730 Homo sapi
955	16	1.1	190	11	AF066731	AF066731 Homo sapi
956	16	1.1	190	11	AF066732	AF066732 Homo sapi
957	16	1.1	190	11	AF066734	AF066734 Homo sapi
958	16	1.1	190	11	AF066735	AF066735 Homo sapi
959	16	1.1	190	11	AF066736	AF066736 Homo sapi
960	16	1.1	190	11	AF066737	AF066737 Homo sapi
961	16	1.1	190	11	AF066738	AF066738 Homo sapi
962	16	1.1	190	11	AF066739	AF066739 Homo sapi
963	16	1.1	190	11	AF066741	AF066741 Homo sapi
964	16	1.1	190	11	AF066742	AF066742 Homo sapi
965	16	1.1	190	11	AF066743	AF066743 Homo sapi
966	16	1.1	190	11	AF066744	AF066744 Homo sapi
967	16	1.1	190	11	AF066745	AF066745 Homo sapi
968	16	1.1	190	11	AF066746	AF066746 Homo sapi
969	16	1.1	190	11	AF066748	AF066748 Homo sapi
970	16	1.1	190	11	AF066749	AF066749 Homo sapi
971	16	1.1	190	11	AF066750	AF066750 Homo sapi
972	16	1.1	190	11	AF066751	AF066751 Homo sapi
973	16	1.1	190	11	AF066752	AF066752 Homo sapi
974	16	1.1	190	11	AF066753	AF066753 Homo sapi
975	16	1.1	190	11	AF066754	AF066754 Homo sapi
976	16	1.1	190	11	AF066755	AF066755 Homo sapi
977	16	1.1	190	11	AF066756	AF066756 Homo sapi
978	16	1.1	190	11	AF066758	AF066758 Homo sapi
979	16	1.1	190	11	AF066759	AF066759 Homo sapi
980	16	1.1	190	11	AF066760	AF066760 Homo sapi
981	16	1.1	190	11	AF066762	AF066762 Homo sapi
982	16	1.1	190	11	AF066763	AF066763 Homo sapi
983	16	1.1	190	11	AF066767	AF066767 Homo sapi
984	16	1.1	190	11	AF066774	AF066774 Homo sapi
985	16	1.1	190	11	AF066776	AF066776 Homo sapi
986	16	1.1	190	11	AF066777	AF066777 Homo sapi
987	16	1.1	190	11	AF066779	AF066779 Homo sapi
988	16	1.1	190	11	AF066782	AF066782 Homo sapi
989	16	1.1	190	11	AF066790	AF066790 Homo sapi
990	16	1.1	190	11	AF066791	AF066791 Homo sapi
991	16	1.1	190	11	AF066792	AF066792 Homo sapi
992	16	1.1	190	11	AF066793	AF066793 Homo sapi
993	16	1.1	190	11	AF066794	AF066794 Homo sapi
994	16	1.1	201	8	PMKCHANN2	PMKCHANN2 Homo sapi
995	16	1.1	207	34	PYTGEBETA	PYTGEBETA Homo sapi
996	16	1.1	210	8	PMKCHANN1	PMKCHANN1 Homo sapi
997	16	1.1	215	8	HVKCHANN	HVKCHANN Homo sapi

998	16	1.1	215	8	SCKCHANN1	Y09752 S.cereale m
999	16	1.1	215	8	VFKCHANN	Y09749 V.faba mRNA
1000	16	1.1	216	8	ZMKCHANN	Y09747 Z.mays mRNA
ALIGNMENTS						
RESULT 1						
PDETFAB/c	PDETFAB	2440 bp	DNA	BCT	27-OCT-1993	
LOCUS	Paracoccus denitrificans	electron transfer flavoprotein alpha and beta subunit genes, complete cds's.				
DEFINITION	LI4864					
ACCESSION	LI4864.1	GI:309664				
VERSION						
KEYWORDS	electron transfer flavoprotein alpha-subunit; electron transfer flavoprotein beta-subunit.					
SOURCE	Paracoccus denitrificans	DNA.				
ORGANISM	Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group; Paracoccus					
REFERENCE	1 (bases 1 to 2440)					
AUTHORS	Bedzyk,L.A., Escudero,K.W., Gill,R.E., Griffin,K.J. and Fireman,F.E.					
TITLE	Cloning, sequencing, and expression of the genes encoding subunits of Paracoccus denitrificans electron transfer flavoprotein					
JOURNAL	J. Biol. Chem. 268 (27), 20211-20217 (1993)					
MEDLINE	93388590					
FEATURES						
source	location/Qualifiers					
	1..2440					
	/organism="Paracoccus denitrificans"					
	/db_xref="taxon:268"					
	411..1169					
	/codon_start=1					
	/transl_table=11					
	/product="electron transfer flavoprotein beta-subunit"					
	/protein_id="AA03071.1"					
	/db_xref="GI:309665"					
	/translation="MKVLVPVKRLIDYNYKARKVSDGVDLANKYMSNPDETAVE EARRLEKQAEETIVASIGYKQAAETLTALAKGDRAILIVAADVQODIEPLAVA KILAAVARAEGTELLIAGKQAIIDNDNATGOMIALIIGAAQTFPAKSEIEGAKKVT REVDGLOTFIVASLPAVVTADRLNLPRAASLPNIKAKKRLDEKTAADYGVDAVPR LEVYSVREPEGRAGIKVGSVDVLGKLEAGVI"					
	1169..2095					
	/codon_start=1					
	/transl_table=11					
	/product="electron transfer flavoprotein alpha-subunit"					
	/protein_id="AA03072.1"					
	/db_xref="GI:309666"					
	/translation="MAVLLIGETNGALNRDATAKAAVAAKALGDVTVLCAGASAKAA AEEAAKIAGVAVLVAEADALYGHRIAEPTAAILVIGAGYSHIAPATPDANKWPRV AALLDPVMSDYSALIDADTPERPPIYAGAAIOVYKSKDKKFTTITASFDAAGEGT APVETETAAADPELSWVADEVAESDRPELTSAARRVSGRGISRESFALIEELADK LGAAVGASRAAVDSGYAPNDQVGGYKGVAAVELYVAVGSIGAIOLHLAGMKDSKIYVA INKDEAPFIQIADYGLVDLFSVPDELGKL"					
	2095..770					
	BASE COUNT	437 a	770 c	869 g	364 t	
ORIGIN						
Query Match	1.3%;	Score 20;	DB 1;	Length 2440;		
Best Local Similarity	100.0%;	Pred. No. 10;				
Matches	20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1315	GCCACGCCCTTCGCCGCGC	1334			
DB	1239	GCCACGCCCTTCGCCGCGC	1220			
RESULT 2						
AFACHRRRA	7390 bp	DNA	BCT	26-MAR-1999		
LOCUS	Ralstonia eutropha Chrb (chrb), Chrc (chrc), Chrd (chrd), Yb1b (yb1b), p1rin, and heat shock protein sigma-32 (RP32)					

genes, complete cds.
 J05278 GI:4521952
 Ralstonia eutropha.
 Ralstonia eutropha
 Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
 Ralstonia.
 1 (bases 1 to 7390)
 Nies,A., Nies,D.H. and Silver,S.
 Nucleotide sequence and expression of a plasmid-encoded chromate
 resistance determinant from Alcaligenes eutrophus
 J. Biol. Chem. 265, 5648-5653 (1990)
 90202806
 2 (bases 1 to 7390)
 Peitzsch,N.
 Transcriptional regulation of chromate metabolism in Ralstonia
 spec. CH34
 Unpublished
 3 (bases 1 to 7390)
 Nies,D.H.
 Direct Submission
 Submitted (26-APR-1993) Ecological and Applied Microbiology,
 Institute for Microbiology, Martin-Luther-University,
 Kurt-Mothes-Str., Halle 3/D-06099, Germany
 4 (bases 2644 to 7390)
 Nies,D.H.
 Direct Submission
 Submitted (26-MAR-1999) Ecological and Applied Microbiology,
 Institute for Microbiology, Martin-Luther-University,
 Kurt-Mothes-Str., Halle 3/D-06099, Germany
 Sequence update by submitter
 On Mar 26, 1999 this sequence version replaced gi:141911.
 Location/Qualifiers
 1. 7390
 /organism="Ralstonia eutropha"
 /strain="CH34"
 /db_xref="taxon:510"
 184. 188
 184. 2415
 RBS
 mRNA
 /note="Inducible by chromate; transcribes both chrA and
 chrB"
 /product="ChrB"
 194. 259
 /gene="chrB"
 194. 784
 /gene="chrB"
 194. 784
 CDS
 /gene="chrB"
 /function="involved in chromate resistance"
 /codon_start=1
 /transl_table=11
 /product_id="AA21963.1"
 /protein_id="GI:141912"
 /db_xref="GI:141912"
 /translation="MNAIPSSPTAMLLVVSPTSGSTARIRHFRGIRKALGATLRD
 GATLPWPGRLAPLOTLATDASDEKVMVSVQADDOQLAEYRALFRTSTYEAEM
 MWELSSARSTLSDSEAEELVARRRGIGIDALRKVDFFPNASARAELOWRDFNMAI
 DIILSPGEPIGVGNIPRRDPTQYOGROWRPASIG"
 260. 781
 /gene="chrB"
 /product="ChrB"
 1202. 1204
 1210. 2415
 /gene="chrA"
 1210. 2415
 /gene="chrA"
 /function="chromate efflux pump"
 /codon_start=1
 /transl_table=11
 /product_id="ChrA"
 /protein_id="AA21964.1"
 /db_xref="GI:141913"
 mat_peptide
 RBS
 gene
 CDS

/translation="MNSPPDPTTAAGSVHTAPTYTLRLQVLMYFLRLGTLGFCGPVAL
 AGYMRDLVEAKOMITPDADYKREGIALAQLAPLAQALAIYLGIVRYRYGATLVYA
 EVLPSPILMTALMAYVRPGLTVMOSVTFGVCAATYIGTIAITAYTLTKRSVNDKLL
 WFTYLVAVATVITSESVAMFLTAGVLFVFMFAPRPMFLMKNFAATPTLPASGM
 MSTLMDPLLSQIGVFPAKAGAFVSGSLAVPLPYGVVTEYHMLNDKQFVDVAAM
 ITPGPVATVIGFYLAVAGLPGACVAAAFPLCYLFTVLPAPYFKYGLPAIILAV
 DGVAAAIAGITGAVIVLAKRSIVDIPITALLVFTALLKFKLSEPMIVAGAILG
 LVAVPLLHH"
 2291. 3337
 /gene="chc"
 /note="Inducible by chromate"
 /product="Chc"
 2291. 3337
 /gene="chc"
 2424. 2428
 /gene="chc"
 2435. 3337
 /gene="chc"
 /function="Mn superoxide dismutase"
 /codon_start=1
 /transl_table=11
 /product="Chc"
 /protein_id="AA21965.2"
 /db_xref="GI:4521953"
 /translation="MLYEMKPLGCEPAKLTLGSEKLIFSHYENNYGGAVRKLNAITAT
 LAFLEMAATAPVPLNLGKREELIATNSMLIHEYPSLGGCSLDKATKAIIRDGCS
 VERKQAEFTMGKALGGSGSWLITTSYPRDGRLVNWDASHATFLAGTIVLADME
 HSYHMDYGAAVAVDAFMQNIHMORANALRCRCRGLTEAGT PMPNATSSLELAAL
 HGNAPDLIDVRRKPAFDASEOMIAGNAMPNDELGMWITLDANRPVTVYCVHGHOY
 SOCCALIEHLILLALARRFSYGMT"
 3257. 4160
 /gene="chrd"
 /note="Inducible by chromate"
 /product="Chrd"
 3257. 4160
 /gene="chrd"
 3343. 3347
 /gene="chrd"
 3354. 4160
 /gene="chrd"
 /function="ATPase subunit of unknown ABC transporter"
 /codon_start=1
 /transl_table=11
 /product="Chrd"
 /protein_id="AAD21772.1"
 /db_xref="GI:4521956"
 /translation="MEKTVSIENVGQFTKTRKGFVALRDIIDLSIAEGEFVTLIGHSG
 CGKSTLNLTAGLTRPTTGALICAGRIAGPRERAVFQNHSLPMTGFEVNHGIV
 ERYGQSEPPAOLKARPHDHLVGLTHARESKYYPHEISGMKORQYIARALMOPKY
 LMDPERGALDALTRAHLDALIKIYARTSTYVWYTHYDEAVLLADRIYMTNGPA
 ATIGELIRVDIPPRDRVTLADDKTFHACTAVLDPLYRKQNPALAEAA"
 4234. 5416
 /gene="ybiB"
 /note="Inducible by chromate"
 /product="ybiB"
 4234. 5416
 /gene="ybiB"
 4251. 4254
 /gene="ybiB"
 4259. 5416
 /gene="ybiB"
 /codon_start=1
 /transl_table=11
 /product="ybiB"
 /protein_id="AAD21774.1"
 /db_xref="GI:4521958"
 /translation="MTSASRAPPAKYIKETIGRGVNGARALPREDAOLFDMALAGR
 VSDLEGAIVLATRIKSEAPHELAGMLDAHDHCAPLAAPADPPVYITPYENKARKOP
 NLVPLAMLLARLDGVPPTLVHGIRFEPGRVSMALFEALGVPLCATQSRDADAMRGIE
 SAGRAADRAVARAVATARQARDHRPQFGAYARODATGRTLPQGIADLDLHPR
 VSGNADRVFLATRRORDAGPGTGGSGRGAPOOPYLARRPSADGRRSRFRFGRC
 ARTAAGHGHRDGRVDSOGAGRGGTGARARCHGGGQGVGAGHAGDGLGQAVTPPA
 TGCRRMARQROGRGSGMSVSPGTCGNMSSQGRMAASSIASLLATATKQ"

```

TITLE
JOURNAL

COMMENT
-----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
-----
Center project name: L3654
Center clone name: 17_L12
-----
Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 118835 bases at least Q40
Consensus quality: 123714 bases at least Q30
Consensus quality: 125172 bases at least Q20
Insert size: 127000; agarose-fp
Insert size: 126105; sum-of-contigs
Quality coverage: 4.8 in Q20 bases; agarose-fp
Quality coverage: 4.9 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
*   1      1266: contig of 1266 bp in length
* *       gap of unknown length
* *   1267      2952: contig of 1686 bp in length
* *       gap of unknown length
* *   2953      5041: contig of 2089 bp in length
* *       gap of unknown length
* *   5042     10634: contig of 5593 bp in length
* *       gap of unknown length
* *   10635    16359: contig of 5725 bp in length
* *       gap of unknown length
* *   16360    22741: contig of 6382 bp in length
* *       gap of unknown length
* *   22742    29444: contig of 6703 bp in length
* *       gap of unknown length
* *   29445    42123: contig of 12679 bp in length
* *       gap of unknown length
* *   42124    55334: contig of 13211 bp in length
* *       gap of unknown length
* *   55335    86293: contig of 30959 bp in length
* *       gap of unknown length
* *   86294   126105: contig of 39812 bp in length.
*         Location/Qualifiers
*           1..126105
*             /organism="Homo sapiens"
*             /db_xref="taxon:9606"
*             /clone="RP11-17L12"
*             /clone_id="RPC1-11 Human Male BAC"
BASE COUNT      39023 a 26923 c 24756 g 35402 t          1 others
ORIGIN
Query Match      1.3%; Score 20; DB 45; Length 126105;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 407 TCGAGGCGAGTCGCCCTG 426
Db 27270 TCGAGGCGAGTCGCCCTG 27289

RESULT 4
LOCUS A94242 128 bp DNA PAT 26-JAN-2000
DEFINITION Sequence 7 from Patent WO9949036.
ACCESSION A94242
VERSION A94242.1 GI:6778881
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.

REFERENCE
AUTHORS Wingate,V.P. and Wolf,M.A.
TITLE LEPIDOPTERAN GABA GATED CHLORIDE CHANNEL AND METHODS OF USE THEREOF
JOURNML Patient: WO 9949036-A-30-SEP-1999.
Rhone Poulenc Inc (US); WINGATE VINCENT PAUL MARY (US)

FEATURES
source
1. .128
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 26 a 34 c 32 g 36 t
ORIGIN

Query Match 1.3%; Score 19; DB 5; Length 128;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 305 TCGAGCGGAGACCATGAC 323
Db 127 TCGAGCGGAGACCATGAC 109

RESULT 5
REJ001389 1878 bp DNA BCT 27-JUL-1998
LOCUS Rhizobium etli ribose kinase gene.
DEFINITION AJ001389
ACCESSION AJ001389.1 GI:3355834
KEYWORDS rbsk gene; ribose kinase.
SOURCE Rhizobium etli.
ORGANISM Rhizobiaceae; Rhizobium.
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Rhizobiaceae; Rhizobium.
1 (bases 1 to 1878)
AUTHORS Michiels,J.
TITLE Direct Submission
JOURNML Submitted (26-JAN-1998) Michiels J., Faculty Agr. Appl. Biol. Sci.,
F.A. Janssens Lab. of Genetics, K.U. Leuven, Kardinaal Mercierlaan
92, Heverlee, 3001, BELGIUM
2 (bases 1 to 1878)
AUTHORS Michiels,J., Verreth,C. and Vanderleyden,J.
TITLE Sequence analysis of the Rhizobium etli ribose kinase gene rbsk and
its phylogenetic position
JOURNML Unpublished
FEATURES
source
1. .1878
/organism="Rhizobium etli"
/strain="CNPAF512"
/db_xref="taxon:29449"
113. .1015
/gene="Rbsk"
113. .1015
/gene="Rbsk"
/function="Ribose kinase"
/codon_start=1
/transl_table=11
/product="Rbsk"

gene
CDS

/protein_id="CA04719.1"
/db_xref="GI:3355835"
/db_xref="SPTREMBL:O86458"
/translation="MTVFSGINDLATTDRLPKPEETVAGNSFATAAGKGAQNAL
AARRAGRYVHMAGAVGKDGFAADALALDADGDLSPVKHVDGPTGALLIVGDGEN
MAAVPGANGILVTADAETAVGKMSGGDILMLQLEVPDVERALSRARVYSILN
LAPLPAPRLGRLLADIVIANETFEERLAGODGDWADAREALHRLHQETLIVTL
GGDVLTAIRNGVISRAGLKEPDPYVAGADPTGCTTASLDEBLRIFASALRAAVA
GSLACLAGAPSPISLPQDVADRI"

BASE COUNT 338 a 607 c 569 g 354 t 10 others
ORIGIN

Query Match 1.3%; Score 19; DB 1; Length 1878;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GGAGAAATGCCCGTGGCG 454
Db 1272 GGAGAAATGCCCGTGGCG 1290

RESULT 6
VARRNSQA/c 1891 bp DNA PLN 30-JUN-1993
LOCUS V.alboatrum large sub-unit rRNA, 3' end and small sub-unit rRNA, 5'
DEFINITION end.
ACCESSION Z23132
VERSION Z23132.1 GI:312961
KEYWORDS large sub-unit rRNA; ribosomal RNA; small sub-unit rRNA.
SOURCE Verticillium albo-atrum.
ORGANISM Verticillium albo-atrum
Eukaryota; Fungi; Ascomycota; Euascomycetes; Pyrenomycetes;
Hypocreales; Hypocreaceae; mitosporic Hypocreaceae; Verticillium.
1 (bases 551 to 878)
AUTHORS Morton,A., Rabrett,A.M., Carder,J.H. and Barbara,D.J.
TITLE The occurrence of sub-repeat sequences in the intergenic regions of
the rRNA genes of Verticillium dahliae and V. albo-atrum
JOURNML Unpublished
REFERENCE 2 (bases 1 to 1891)
AUTHORS Barbara,D.J.
TITLE Direct Submission
JOURNML Submitted (28-JUN-1993) Derek J. Barbara, Plant Pathology and Weed
Science, Horticulture, Research International, East Malling, West
Malling, Kent, ME19, 6BU, United Kingdom
FEATURES
source
1. .1891
/organism="Verticillium albo-atrum"
/isolate="1974"
/db_xref="taxon:27335"
/clone="pVAL12(pVR94)"
/clone_lib="pVARRNA, pUC18 library of A Morton."
1. .218
/product="large sub-unit rRNA"
219. .1887
rna repeat_region 551. .878
rna 1888. .>1891
/product="small sub-unit rRNA"
BASE COUNT 405 a 466 c 584 g 409 t 27 others
ORIGIN

Query Match 1.3%; Score 19; DB 7; Length 1891;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 297 TGAGGAATCGAGCGGAG 315
Db 221 TGAGGAATCGAGCGGAG 203

RESULT 7
CGU16848/c

LOCUS GCUI6848 5026 bp mRNA VRT 04-APR-1995
DEFINITION Gallus gallus complement C3 precursor mRNA, complete cds.
ACCESSION U16848
VERSION U16848.1 GI:755814
KEYWORDS
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE
AUTHORS Mavroidis, M., Sunyer, J.O. and Lambiris, J.D.
TITLE Isolation, primary structure, and evolution of the third component
of chicken complement and evidence for a new member of the alpha
2-macroglobulin family
J. Immunol. 154 (5), 2164-2174 (1995)
MEDLINE 95173423
REFERENCE 2 (bases 1 to 5026)
AUTHORS Lambiris, J.D.
TITLE Direct Submission
JOURNAL Submitted (03-NOV-1994) John D. Lambiris, Pathology, University of
Pennsylvania, 36th St. and Hamilton Walk, Philadelphia, PA
19104-6079, USA
FEATURES
source Location/Qualifiers
1..5026
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="Ch 12, 4.1.1, 24.5.1, 34.1.1.1, 35.1.1.1"
/clone_1b="chicken liver 5'-stretch cDNA library
(Clontech CL1018a)"
/sex="female"
/issue_type="liver"
22..4980
/codon_start=1
/product="complement C3 precursor"
/protein_id="AA64694.1"
/db_xref="GI:755815"
/translation="MGILLPLLLGVLLHAHPPTPAOMTMTPAVLRLLDTDEKVVLE
APGLAATEANILIVDPQPKRYLFQVRKOLNPAEGMMAATVAVPKVLPPVGHKF
VSVARVAGVATLEKVLVLSGSHIFLQDKPITPTGPTVLSLPLASHMOPMLKTY
IVKPKPDNYITIKQPVSSPMRGISFINNIPENVISIGTWITAFEDSDQDFSTQ
FEVKEVYLPSEFTLDPOKEFLIDPAEDRVTTRIKLGNLQGAFAVFEVVAD
EKTIPOSLSORVAVTDGQAVLPMLLRQFANLQELVGHSLYVTVLFEVSQDMY
EAQRSGRIATVSPYTHFTHTPKYFKPMFEDTVTVTNDSNPAAGIPVKADNFG
LVSTORDGATKLVLNMPANKNSVITVTRQDKLPERORSROIIVAEYOSGNSGY
LHLAVASQVOPGDNLPINFLKSNRDDVRSVSTFYLLISGLIYVHGRORBDQ
SLVTMLSPYANLIPSPRIYAYHNGELIASVWVDKDTGMSLVYRGASEADNR
VHEPRTPMKRHIEGDKAHAGLVAVKAVYLVKNNLQSKWVDIVENSQIGCTPSG
RNQGVAFADAGLSLTSNVNINTEORSEVOCAPAKRRRSVRLLIKHGTMAEDSKN
LRKQCDGIRKNNLMGYSCEKRAYVLDAKSCAPFLSCCLYIGIDEEDEELQELAR
SEVDDAFLSDEDITSRLSPESLMQVELETPNQGISMKPLIYLNKDSITTYVL
AVSISNKGICVADPYEITVMKEFFIDRLPYSAVNEOVEVRAILLNWTNKKIKRY
ELMNPALCSASTTKTRTQOOLIPOLPOSSDAVFPVPLETGHNDVYKRAVNSYS
DGKKRLRYPEBGMLEKTYKIVELDPKLGNGGVEVKKANLSDIVPSESEKRV
SLOGNPVSLIVERKATDGTKLKLIVTPSGGEGNMIGMTPTVAVHLDSTMTEG
INRTEAIELIKKGYTOOLAYRKEDGSFAFTTRPSTTAVAVAFMAIMVQIK
PEVVCAGAILKMLLEKQDPGLFQEDAPVIHKEWGVHGAEPVSLTAFAVLAJLOSQ
KICKNVKSLDGSIAKASDYLSRKYOSLRPYVALSYALATGKLSKRVLMKRSK
DGHMAERNAHYNITGTSYALVALLOMERKELTGYVRRLAQQNFGGSGYGTQATI
LVQALAOYHVALPRVHMLNDVSLPRKANITRIENNNALVARSATKLNEDPT
VRAEGTGKGTMTVTVYKAKVPERKNCNFDLRVSEVDEKAGREVEGYIRSVKLTIC
TRFLTDVATMTSLIDISMLTAFSPVODKLSSEGVERYISKFEIDHALSNRSLIY
LDVYSHQVEECIAFRAHQHFOVGLIOPASIVSYKIDRCRTFRYHDPKAGDLKRI
CHGEVCAEENCIRAYKNDNPITVNERDIACKRGVDYVYKVVAVNTEEPKSHDNTIM
SLITVTKMGTDEMPGSSNRPFSHKCRDALSLQKQDYLVMGLASDLMTVTSRSTYL
ISKDVTLEAMPLEBSCQDADLPDLCDFTEFSNVLVFGCP1"
slg_peptide
mat_peptide
mat_peptide
mat_peptide
BASE COUNT 1157 a 1520 c 1444 g 905 t
ORIGIN

Query Match 1.3%; Score 19; DB 4; Length 5026;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 88 CAGTTGGAAGATCTGCTGG 106
|||||
Db 2661 CAGTTGGAAGATCTGCTGG 2643
RESULT 8
LOCUS 039714
DEFINITION Mycoplasma genitalium section 36 of 51 of the complete genome.
ACCESSION U39714 L43967
VERSION U39714.1 GI:3844905
KEYWORDS
SOURCE
ORGANISM Mycoplasma genitalium.
Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
Mycoplasmataceae; Mycoplasma.
REFERENCE
AUTHORS Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A.,
Fleischmann, R.D., Bult, C.J., Kerlavage, A.R., Sutton, G.,
Kiley, J.M., Fritchman, J.L., Weidman, J.F., Small, K.V., Sandusky, M.,
Fuhrmann, J.L., Nguyen, D.T., Uterback, T., Saudak, D.M.,
Phillips, C.A., Merrick, J.M., Tomb, J., Dougherty, B.A., Bott, K.F.,
Hu, P.C., Lucier, T.S., Peterson, S.N., Smith, H.O. and Venter, J.C.
The minimal gene complement of Mycoplasma genitalium
Science 270 (5235), 397-403 (1995)
MEDLINE 96026346
REFERENCE 2 (bases 1 to 12486)
AUTHORS Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A.,
Fleischmann, R.D., Bult, C.J., Kerlavage, A.R., Sutton, G.,
Kiley, J.M., Fritchman, J.L., Weidman, J.F., Small, K.V.,
Sandusky, M., Fuhrmann, J.L., Nguyen, D.T., Uterback, T.R.,
Saudak, D.M., Phillips, C.A., Merrick, J.M., Tomb, J.-F.,
Dougherty, B.A., Bott, K.F., Hu, P.-C., Lucier, T.S., Peterson, S.N.,
Smith, H.O., Hutchinson, C.A. III. and Venter, J.C.
Direct Submission
Submitted (29-OCT-1995) The Institute for Genomic Research, 9712
Medical Center Drive, Rockville, MD 20850, USA
3 (bases 1 to 12486)
REFERENCE
AUTHORS Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A.,
Kiley, J.M., Fritchman, J.L., Weidman, J.F., Small, K.V., Sandusky, M.,
Fuhrmann, J.L., Nguyen, D.T., Uterback, T., Saudak, D.M.,
Phillips, C.A., Merrick, J.M., Tomb, J., Dougherty, B.A., Bott, K.F.,
Hu, P.C., Lucier, T.S., Peterson, S.N., Smith, H.O. and Venter, J.C.
Direct Submission
Submitted (19-OCT-1998) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
On Nov 5, 1998 this sequence version replaced gi:1046022.
Location/Qualifiers
1..12486
/organism="Mycoplasma genitalium"
/isolate="G37"
/db_xref="taxon:2097"
complement(14..901)
/gene="MG326"
complement(14..901)
/gene="MG326"
/note="similar to GB:U00089 SP:P75312 PID:1674048 percent
identity: 68.94; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=4
/product="conserved hypothetical protein"
/protein_id="AACT71550.1"
/db_xref="GI:3844906"
/translation="MKRTAIIIDTSASIKRGEINGVYILPQVYDGEKSFSDGIEID
YDHYHKLKENPHGNLSTSLPQSDLLKIFEIKRYRFLFLPSKLSGLSTYDMLV
QLAKELSEQNKDEFLVFETSDIAISLKLVLVDIKALVDKGDQGTIKAKVESHKONI

Query Match 1.3%; Score 19; DB 2; Length 12486;
Best local Similarity 100.0%; Pred. No. 38;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

793 ATATTAGCACTTGTCGA 811


```
repeat_region 1072. .1198
/note="MIR repeat: matches 13. .143 of consensus"
repeat_region 1205. .1511
/note="AluSx repeat: matches 1. .308 of consensus"
repeat_region 1751. .1841
/note="MER63 repeat: matches 26. .752 of consensus"
repeat_region 1891. .1931
/note="L2 repeat: matches 2664. .2704 of consensus"
misc_feature 2297. .2644
/note="match: GSS: Em:AQ218889"
repeat_region 2514. .2569
/note="MIR repeat: matches 113. .170 of consensus"
repeat_region 2592. .2887
/note="AluSc repeat: matches 1. .298 of consensus"
repeat_region 3001. .3085
/note="L2 repeat: matches 2619. .2710 of consensus"
repeat_region 3088. .3127
/note="5 copies 8 mer atgaatga 92 conserved"
repeat_region 3839. .4037
/note="MIR repeat: matches 31. .248 of consensus"
repeat_region 4197. .4326
/note="MER5A repeat: matches 20. .149 of consensus"
repeat_region 4343. .4626
/note="Alu repeat: matches 25. .305 of consensus"
repeat_region 5079. .5125
/note="L2 repeat: matches 2686. .2735 of consensus"
repeat_region 5126. .5429
/note="Alu repeat: matches 1. .303 of consensus"
repeat_region 5439. .5743
/note="Alu repeat: matches 1. .300 of consensus"
repeat_region 6365. .6649
/note="AluSx repeat: matches 27. .311 of consensus"
repeat_region 7636. .7708
/note="L2 repeat: matches 2638. .2710 of consensus"
repeat_region 8325. .8386
/note="31 copies 2 mer tt 72 conserved"
repeat_region 8328. .8387
/note="12 copies 5 mer tcttt 86 conserved"
repeat_region 8412. .8493
/note="Alu repeat: matches 225. .308 of consensus"
repeat_region 9387. .9483
/note="L2 repeat: matches 2611. .2710 of consensus"
repeat_region 10720. .11321
/note="L2 repeat: matches 2065. .2709 of consensus"
repeat_region 11479. .11599
/note="MIR repeat: matches 31. .145 of consensus"
repeat_region 12182. .12267
/note="43 copies 2 mer cc 64 conserved"
repeat_region 13334. .13419
/note="MIR repeat: matches 44. .144 of consensus"
repeat_region 13871. .14283
/note="MIR repeat: matches 1. .391 of consensus"
repeat_region 15818. .15839
/note="match: GSS: Em:AQ314743; match: STS: Em:G56767"
repeat_region 15818. .15839
/note="11 copies 2 mer tt 100 conserved"
misc_feature 15818. .15839
/note="match: GSS: Em:AQ740632"
misc_feature 16460. .16520
/note="match: GSS: Em:AQ740620"
repeat_region 16460. .16520
/note="MER3 repeat: matches 146. .209 of consensus"
misc_feature 16460. .16520
/note="match: GSS: Em:AQ739191"
repeat_region 16521. .16647
/note="FLAM_A repeat: matches 1. .133 of consensus"
repeat_region 17352. .17478
/note="MIR repeat: matches 32. .148 of consensus"
repeat_region 18248. .18455
/note="MIR repeat: matches 51. .262 of consensus"
repeat_region 18484. .18655
/note="MIR repeat: matches 84. .256 of consensus"
repeat_region 19753. .20050
```

```
repeat_region /note="Alu repeat: matches 1. .293 of consensus"
20336. .20350
repeat_region /note="MIR repeat: matches 246. .259 of consensus"
20351. .20410
repeat_region /note="L2 repeat: matches 2666. .2729 of consensus"
20411. .20476
repeat_region /note="MIR repeat: matches 40. .246 of consensus"
20554. .20642
repeat_region /note="L2 repeat: matches 2531. .2620 of consensus"
20687. .20747
repeat_region /note="MIR repeat: matches 93. .153 of consensus"
21468. .22247
repeat_region /note="L2 repeat: matches 1833. .2750 of consensus"
22466. .22617
repeat_region /note="match: GSS: Em:AQ594604"
22640. .22977
repeat_region /note="MIR repeat: matches 82. .262 of consensus"
23096. .23274
repeat_region /note="MIR repeat: matches 10. .192 of consensus"
23902. .24207
repeat_region /note="Alu repeat: matches 1. .312 of consensus"
24743. .24836
repeat_region /note="MIR repeat: matches 35. .142 of consensus"
24996. .25049
repeat_region /note="MIR repeat: matches 92. .148 of consensus"
25050. .25361
repeat_region /note="Alu repeat: matches 1. .312 of consensus"
25410. .25700
repeat_region /note="AluSx repeat: matches 1. .291 of consensus"
26006. .26197
repeat_region /note="L2 repeat: matches 7546. .7734 of consensus"
26198. .26512
repeat_region /note="AluSx repeat: matches 1. .312 of consensus"
26520. .32881
misc_feature /gene="MGAT3"
26520. .32881
repeat_region /note="match: cDNAs: Em:E11285; match: ESTs: Em:A1452685
Em:A1524129 Em:A1198445 Em:A1439025 Em:A1075278
Em:AA743691 Em:AA421504 Em:A1090769 Em:A1308859
Em:AA829826 Em:A1571173 Em:A4501937 Em:AA421473
Em:AA360841 Em:H19142 Em:H29553 Em:AA345131 Em:H47026
Em:A1572262 Em:AA876086 Em:A164862 Em:A1367012 Em:H92616
Em:H41010"
26520. .32881
repeat_region /evidence=not_experimental
26520. .32881
repeat_region /product="dJ1104E15.1 (mannosyl (beta-1,4-)-glycoprotein
beta-1,4-N-acetylglucosaminyltransferase)"
26834. .27125
repeat_region /gene="MGAT3"
26834. .27125
repeat_region /note="AluSx repeat: matches 7. .298 of consensus"
27188. .27489
repeat_region /note="Alu repeat: matches 1. .302 of consensus"
28035. .30123
misc_feature /gene="MGAT3"
```

```
Query Match 1.3%; Score 19; DB 11; Length 112460;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 636 ATCTGTGATCTGTGGTG 654
|||||
Db 101533 ATCTGTGATCTGTGGTG 101551

RESULT 11
LOCUS AC011325/c DNA 145873 bp 19-FEB-2000
DEFINITION Homo sapiens clone Rp11-17P4, *** SEQUENCING IN PROGRESS ***, 44
ACCESSION AC011325
VERSION AC011325.4 GI:7007739
KEYWORDS HIG; HTGS_PHASE1.
```

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 145873)
AUTHORS Muzny,D.M., Adams,C., Bailey,M., Barbara,J., Blankenburg,K.,
Bodocia,B., Bouck,D., Bowle,S., Brooks,A., Bunay,C., Bunick,C.,
Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
Guevara,M., Harris,K., Hernandez,J., Hodgson,A., Hognes,M.,
Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
Licharge,O., Liu,J., Liu,M., Logan,O., Lozado,R.J., Lu,J.,
Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mel,G., Morgan,M.,
Morris,S., Nash,S., Nelson,A., Nguyen,N., Nguyen,S.,
Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
Oulles,M., Reller,D., Rives,M., Samuel,S., Say,D., Scherer,S.,
Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sugeng,R.,
Tabor,P., Taylor,T., Vasquez,L., Vinson,R., Vo,O., Wahbah,M.,
Wattlington,S., Weinstein,G., Weinstein,I.R., Williamson,A.,
Worley,K., Wren,J., Wrenford,G., Yu,W., Zhou,X., Nelson,D. and
Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 145873)
Worley,K.C.
Direct Submission
Submitted (06-OCT-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Feb 19, 2000 this sequence version replaced g1:6560805.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: RP11-17P4
Center clone name: RP11-17P4
----- Summary Statistics
Sequencing vector: M13; L08821
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.980611
Consensus quality: 100197 bases at least Q40
Consensus quality: 115249 bases at least Q30
Consensus quality: 123956 bases at least Q20
Estimated insert size: 145873; agarose-fp estimation
Estimated insert size: 139757; sum-of-ctnigs estimation
Quality coverage: 1.8x in Q20 bases; agarose-fp estimation
Quality coverage: 1.8x in Q20 bases; sum-of-ctnigs estimation

* NOTE: This is a 'working draft' sequence. It currently
* consists of 44 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 847: contig of 847 bp in length
* 868 867: gap of unknown length
* 2011 2010: contig of 1143 bp in length
* 2031 2873: gap of unknown length
* 2874 2893: gap of 843 bp in length
* 2894 4043: contig of 1150 bp in length
* 4044 4063: gap of unknown length
* 4064 5087: contig of 1024 bp in length
* 5088 5107: gap of unknown length
* 5108 6154: contig of 1047 bp in length
* 6155 6174: gap of unknown length

6175 6997: contig of 823 bp in length
6998 7017: gap of unknown length
7018 8000: contig of 983 bp in length
8001 8020: gap of unknown length
8021 8792: contig of 771 bp in length
8792 8811: gap of unknown length
8811 9973: contig of 1162 bp in length
9973 9974: gap of unknown length
9974 9994: contig of 1247 bp in length
9994 11240: gap of unknown length
11240 11261: contig of 1105 bp in length
11261 12365: gap of unknown length
12365 12386: contig of 1730 bp in length
12386 14115: gap of unknown length
14115 14136: contig of 1419 bp in length
14136 15554: gap of unknown length
15554 15575: gap of unknown length
15575 16905: contig of 1331 bp in length
16905 16906: gap of unknown length
16906 16926: contig of 1627 bp in length
16926 18553: gap of unknown length
18553 18572: gap of unknown length
18572 19523: contig of 951 bp in length
19523 19524: gap of unknown length
19524 19544: contig of 1460 bp in length
19544 21003: gap of unknown length
21003 21023: gap of unknown length
21023 23362: contig of 2339 bp in length
23362 23363: gap of unknown length
23363 24938: contig of 1556 bp in length
24938 24958: gap of unknown length
24958 26872: contig of 1914 bp in length
26872 26892: gap of unknown length
26892 28533: contig of 1641 bp in length
28533 28553: gap of unknown length
28553 30648: contig of 2095 bp in length
30648 30669: gap of unknown length
30669 32952: contig of 2284 bp in length
32952 32972: gap of unknown length
32972 35237: contig of 2265 bp in length
35237 35257: gap of unknown length
35257 37173: contig of 1916 bp in length
37173 37194: gap of unknown length
37194 38895: contig of 1702 bp in length
38895 38915: gap of unknown length
38915 41633: contig of 2718 bp in length
41633 41653: gap of unknown length
41653 44466: contig of 2813 bp in length
44466 44486: gap of unknown length
44486 46750: contig of 2264 bp in length
46750 46770: gap of unknown length
46770 49825: contig of 3055 bp in length
49825 49845: gap of unknown length
49845 52368: contig of 2523 bp in length
52368 52388: gap of unknown length
52388 54857: contig of 2469 bp in length
54857 54877: gap of unknown length
54877 59298: contig of 4421 bp in length
59298 59318: gap of unknown length
59318 63006: contig of 3688 bp in length
63006 63026: gap of unknown length
63026 66219: contig of 3193 bp in length
66219 66239: gap of unknown length
66239 72271: contig of 6032 bp in length
72271 72291: gap of unknown length
72291 78268: contig of 5977 bp in length
78268 78288: gap of unknown length
78288 83796: contig of 5508 bp in length
83796 83816: gap of unknown length
83816 92717: contig of 8900 bp in length
92717 92736: gap of unknown length
92736 103847: contig of 11111 bp in length
103847 103867: gap of unknown length
103867 117291: contig of 13424 bp in length
117291 117311: gap of unknown length
117311 129740: contig of 12429 bp in length

* 129741 129760: gap of unknown length
* 129761 145873: contig of 16113 bp in length.
Location/Qualifiers
1. 145873
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-17P4"
BASE COUNT 38679 a 33326 c 32779 g 39341 t 1748 others
ORIGIN

Query Match 1.3%; Score 19; DB 57; Length 145873;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1435 AAGACACGCTCAGTAT 1453
|||||
Db 19997 AAGACACGCTCAGTAT 19979

RESULT 12
AP000803 LOCUS
DEFINITION Homo sapiens chromosome 11 clone RP11-642F7 map 11q13, WORKING.
ACCESSION AP000803
VERSION AP000803.1 GI:6997640
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens DNA, clone:RP11-642F7.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 154700)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens 154,700 genomic DNA of 11q13
JOURNAL Published Only in Database (1999) In press
REFERENCE 2 (bases 1 to 154700)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (30-NOV-1999) to the DDBJ/EMBL/Genbank databases.
REFERENCE Mesahra Hattori, The Institute of Physical and Chemical Research
(RIKEN), Genomic Sciences Center (GSC), Kitasato Univ., 1-15-1
Kitasato, Sagamihara, Kanagawa 228-8555, Japan
(E-mail:hattori@gscc.riken.go.jp, URL:http://hgp.gscc.riken.go.jp/,
Tel:81-42-778-9923, Fax:81-42-778-9924)
On Feb 18, 2000 this sequence version replaced gi:6197950.
COMMENT ----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gscc.riken.go.jp/
Contact: hattori@gscc.riken.go.jp
----- Project Information
Center project name: HumDRAFT11
Center clone name: RP11-642F7
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 117360 bases at least Q40
Consensus quality: 127920 bases at least Q30
Consensus quality: 133592 bases at least Q20
Insert size: 136324; sum-of-contigs
Quality coverage: 4.27x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
38 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs 'N', but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved

1 13642 contig of 13642 bp in length
14143 21460 contig of 21460 bp in length
21961 30993 contig of 9033 bp in length
31494 39792 contig of 8297 bp in length
40293 47612 contig of 7320 bp in length
48113 55102 contig of 6990 bp in length
55603 60952 contig of 5350 bp in length
61455 67314 contig of 5860 bp in length
67315 67816 contig of 502 bp in length
67817 73847 contig of 6031 bp in length
73848 74351 contig of 504 bp in length
74352 79192 contig of 4841 bp in length
79193 79695 contig of 503 bp in length
79696 84618 contig of 4923 bp in length
84619 85119 contig of 501 bp in length
85120 88409 contig of 3291 bp in length
88410 88912 contig of 503 bp in length
88913 92062 contig of 3150 bp in length
92063 92563 contig of 500 bp in length
92563 97033 contig of 4471 bp in length

* NOTE: This is a 'working draft' sequence. It currently
* consists of 38 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 13642: contig of 13642 bp in length
* 13643 14144: gap of 502 bp
* 14145 21460: contig of 7316 bp in length
* 21461 21960: gap of 500 bp
* 21961 30993: contig of 9033 bp in length
* 30994 31495: gap of 502 bp
* 31496 39792: contig of 8297 bp in length
* 39793 40293: gap of 500 bp
* 40293 47610: contig of 7318 bp in length
* 47611 48112: gap of 502 bp
* 48113 55102: contig of 6990 bp in length
* 55103 55602: gap of 500 bp
* 55603 60951: contig of 5349 bp in length
* 60952 61454: gap of 503 bp
* 61455 67314: contig of 5860 bp in length
* 67315 67816: gap of 502 bp
* 67817 73847: contig of 6031 bp in length
* 73848 74351: gap of 504 bp
* 74352 79192: contig of 4841 bp in length
* 79193 79695: gap of 503 bp
* 79696 84618: contig of 4923 bp in length
* 84619 85119: gap of 501 bp
* 85120 88409: contig of 3290 bp in length
* 88410 88912: gap of 503 bp
* 88913 92062: contig of 3150 bp in length
* 92063 92563: gap of 500 bp
* 92563 97033: contig of 4471 bp in length

```
* 97034 97534: gap of 501 bp
* 97535 100551: contig of 3017 bp in length
* 100552 101056: gap of 505 bp
* 101057 105058: contig of 4002 bp in length
* 105059 105558: gap of 500 bp
* 105559 108879: contig of 3321 bp in length
* 108880 109385: gap of 506 bp
* 109386 111010: contig of 1625 bp in length
* 111011 111513: gap of 503 bp
* 111514 113186: contig of 1673 bp in length
* 113187 113686: gap of 500 bp
* 113687 116280: contig of 2594 bp in length
* 116281 116782: gap of 502 bp
* 116783 118635: contig of 1853 bp in length
* 118636 119135: gap of 500 bp
* 119136 122093: contig of 2958 bp in length
* 122094 122593: gap of 500 bp
* 122594 125851: contig of 3258 bp in length
* 125852 126352: gap of 501 bp
* 126353 128860: contig of 2508 bp in length
* 128861 129364: gap of 504 bp
* 129365 131579: contig of 2215 bp in length
* 131580 132081: gap of 502 bp
* 132082 133110: contig of 1029 bp in length
* 133111 133612: gap of 502 bp
* 133613 135020: contig of 1408 bp in length
* 135021 135524: gap of 504 bp
* 135525 136942: contig of 1418 bp in length
* 136943 137443: gap of 501 bp
* 137444 138628: contig of 1185 bp in length
* 138629 139135: gap of 507 bp
* 139136 141042: contig of 1907 bp in length
* 141043 141544: gap of 502 bp
* 141545 142669: contig of 1125 bp in length
* 142670 143173: gap of 504 bp
* 143174 144589: contig of 1416 bp in length
* 144590 145093: gap of 504 bp
* 145094 146631: contig of 1538 bp in length
* 146632 147132: gap of 501 bp
* 147133 148188: contig of 1056 bp in length
* 148189 148690: gap of 502 bp
* 148691 149842: contig of 1152 bp in length
* 149843 150342: gap of 500 bp
* 150343 151542: contig of 1200 bp in length
* 151543 152042: gap of 500 bp
* 152043 153137: contig of 1095 bp in length
* 153138 153639: gap of 502 bp
* 153640 154700: contig of 1061 bp in length.
```

FEATURES
source
1. .154700
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-642F7"
/map="11q13"

BASE COUNT 30140 a 37119 c 36402 g 29931 t 21108 others
ORIGIN

Query Match 1.3%; Score 19; DB 32; Length 154700;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1287 CCATGAAGCCTCTCATC 1305
|||||
Db 85837 CCATGAAGCCTCTCATC 85855

RESULT 13
AC020344 157559 bp DNA HTG 03-JAN-2000
LOCUS AC020344 157559 bp DNA HTG 03-JAN-2000
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
pieces.

ACCESSION AC020344
VERSION AC020344.1 GI:6664553
KEYWORDS HTG: HTGS_PHASE2.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 157559)
AUTHORS Adams, M. and Venter, J.C.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT This sequence was identified as CDN:10212849 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1. .157559
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

BASE COUNT 43649 a 34299 c 35451 g 44160 t
ORIGIN

Query Match 1.3%; Score 19; DB 44; Length 157559;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 591 GGAGATGCGCCAGACCTG 609
|||||
Db 142763 GGAGATGCGCCAGACCTG 142781

RESULT 14
AF033522 208 bp DNA STS 24-MAR-1999
LOCUS AF033522/c
DEFINITION Hyla regilla repeat region clone Hyre46, sequence tagged site.
ACCESSION AF033522
VERSION AF033522.1 GI:2645829
KEYWORDS STS.
SOURCE Hyla regilla.
ORGANISM Hyla regilla
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae; Hyla.

REFERENCE 1 (bases 1 to 208)
AUTHORS Call, D.R. and Hallett, J.G.
TITLE PCR primers for microsatellite loci in the anurans Rana
luteiventris and Hyla regilla
JOURNAL Mol. Ecol. 7 (8), 1085-1087 (1998)

JOURNAL 98377747
MEDLINE 2 (bases 1 to 208)
REFERENCE Call, D.R. and Hallett, J.G.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (09-NOV-1997) Zoology, Washington State University,
Pullman, WA 99164-4236, USA

FEATURES
source
1. .208
/organism="Hyla regilla"
/db_xref="taxon:47562"
/clone="Hyre46"

repeat_region 90..127
/rpt_unit=CCAT
/rpt_type=tandem
BASE COUNT 62 a 77 c 25 g 44 t
ORIGIN

Query Match 1.2%; Score 18; DB 13; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 636 ATCTGTGATCTGTGGT 653
|||||
Db 85 ATCTGTGATCTGTGGT 68

RESULT 15

G51319 349 bp DNA STS 25-JUN-1999
LOCUS SHGC-82833 Human Homo sapiens STS genomic, sequence tagged site.
DEFINITION G51319
ACCESSION G51319
VERSION G51319.1 GI:5222646
KEYWORDS STS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 349)
AUTHORS Myers, R.M.
TITLE Human STSs (1999)
JOURNAL Unpublished (1999)
COMMENT
Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myerseshgc.stanford.edu
Primer A: TCCTGCAATAAATGCTTTTC
Primer B: GACTGGAGCTCAGCTTTAAT
STS size: 334

PCR Profile:

Initial incubation: 95 degrees C for 10 minutes
Denaturation: 94 degrees C for 30 seconds
Annealing: 60 degrees C for 30 seconds
Polymerization: 72 degrees C for 23 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9700

Protocol:

Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
AmpliTaq Gold Polymerase: 0.07 units/ul
Total Vol: 5 ul

Buffer:

MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3

FEATURES
source BAC ends sequenced at TIGR from the RPC11 BAC library. Designed
and developed at the Stanford Human Genome Center.
Location/Qualifiers
1..349

/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="5"
/clone_lib="Human"

STS
primer_bind 7..31
primer_bind complement(318..340)
BASE COUNT 79 a 74 c 59 g 137 t
ORIGIN

Query Match 1.2%; Score 18; DB 13; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1287 CCATGAGCCTTCTGTAT 1304
|||||

Db 167 CCATGAGCCTTCTGTAT 184

Search completed: May 11, 2000, 00:41:07
Job time: 8098 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2000, 22:28:18 ; Search time 60.27 Seconds
(Without alignments) 6189,421 Million cpi updates/sec

Title: US-09-080-127-1

Perfect score: 1491
Sequence: 1 ATGAGTCGCTTTGTGGGC.....ATTCTCAAGTCGAGCATAG 1491

Scoring table: OLIGO.MDC
Gapop 60.0 , Gapext 60.0

Searched: 311585 seqs, 125096042 residues

Word size : 0

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1491	100.0	1491	1 X00070	Aspergillus oryzae
2	1491	100.0	1491	1 X82521	Aspergillus oryzae
3	1491	100.0	1491	1 X00018	Aspergillus oryzae
4	21	1.4	32	1 X00024	Aspergillus oryzae
5	21	1.4	32	1 X00012	Aspergillus oryzae
6	20	1.3	23	1 X00021	Aspergillus oryzae
7	20	1.3	23	1 X82525	Aspergillus oryzae
8	20	1.3	32	1 X00023	Aspergillus oryzae
9	20	1.3	32	1 X00011	Aspergillus oryzae
10	19	1.3	128	1 V55014	Fragment of GABA g
11	19	1.3	110000	1 T58840_4	Continuation (5 of
12	18	1.2	4372	1 T91360	Of virus genomic
13	17	1.1	444	1 X40591	Human secreted pro
14	17	1.1	1231	1 T95755	Arabidopsis SCAREC
15	17	1.1	1507	1 002845	Complete sequence
16	17	1.1	1509	1 049057	Bile acid sulphate
17	17	1.1	2253	1 063221	Human masking prot
18	17	1.1	2256	1 063220	Human masking prot
19	17	1.1	2268	1 063219	Human masking prot
20	17	1.1	2535	1 063222	Human masking prot
21	17	1.1	4901	1 X20621	Polynucleotide seq
22	17	1.1	4972	1 Q14345	TF Beta 1 binding
23	17	1.1	5261	1 X02985	Human IL-1ra BAC c
24	17	1.1	10497	1 X03048	Human IL-1ra BAC c
25	16	1.1	127	1 V76088	Staphylococcus aur
26	16	1.1	225	1 N50093	Sequence encoding
27	16	1.1	278	1 V17745	Human breast cance
28	16	1.1	283	1 T63570	Cellulomonas flm1
29	16	1.1	324	1 Q61130	Human brain expres
30	16	1.1	356	1 X40395	Human secreted pro
31	16	1.1	532	1 X20985	Polynucleotide seq
32	16	1.1	590	1 T78381	Cherry polyphenol
33	16	1.1	1150	1 X12973	Enterococcus faeca

34	16	1.1	1200	1 N40302	Phasolin gene ins
35	16	1.1	1236	1 Q79911	Aromatic aldehydri
36	16	1.1	1425	1 T77036	DNA encoding heat-
37	16	1.1	1455	1 T12315	Replicable and tra
38	16	1.1	1524	1 Q10499	Beta-ketothiolase
39	16	1.1	1525	1 N91207	Beta-ketothiolase
40	16	1.1	1548	1 X20523	Polynucleotide seq
41	16	1.1	1734	1 N70995	Sequence encoding
42	16	1.1	1734	1 Q73450	DNA encoding q1 of
43	16	1.1	1767	1 Q06900	Neo-pullulanase ge
44	16	1.1	1842	1 Q88134	Human P20 receptor
45	16	1.1	1972	1 Q10332	ECOR1-BamHI fragme
46	16	1.1	2083	1 V02472	Bacillus amyloliqu
47	16	1.1	2084	1 Q88067	Bacillus amyloliqu
48	16	1.1	2084	1 Q95032	Bacillus amyloliqu
49	16	1.1	2122	1 V54981	Class A potato sta
50	16	1.1	2122	1 V55046	Potato class A sta
51	16	1.1	2259	1 X18189	C. trachomatis gid
52	16	1.1	2259	1 X07334	Streptococcus pneu
53	16	1.1	2364	1 V99969	Nucleotide sequenc
54	16	1.1	2550	1 T47532	Cytomegalovirus To
55	16	1.1	2550	1 T47534	Cytomegalovirus To
56	16	1.1	2724	1 Q78166	Human cytomagalovi
57	16	1.1	2724	1 T47527	Human cytomagalovi
58	16	1.1	2724	1 T91280	Human cytomagalovi
59	16	1.1	2799	1 V20871	Murine phospholipa
60	16	1.1	2945	1 Q87790	S. lavendulae mcr
61	16	1.1	2972	1 V29125	Homo sapiens Smoot
62	16	1.1	3232	1 T12316	Replicable and tra
63	16	1.1	3374	1 V20872	Murine phospholipa
64	16	1.1	3425	1 Q99778	Human phosphatidy
65	16	1.1	3627	1 Q27466	P772 promoter and
66	16	1.1	3627	1 Q53879	Sequence comprisin
67	16	1.1	3830	1 N70938	DNA fragment contg
68	16	1.1	4166	1 N90528	DNA sequence encod
69	16	1.1	4260	1 T47528	H6 promoted cyto
70	16	1.1	4594	1 T47532	CMV gB transmembra
71	16	1.1	4594	1 T47535	CMV gB transmembra
72	16	1.1	4768	1 T47531	H6 promoted cyto
73	16	1.1	4875	1 T96058	E. coli colonisati
74	16	1.1	5234	1 T47563	Cytomegalovirus gB
75	16	1.1	5661	1 V21518	Rattus norvegicus
76	16	1.1	5688	1 V19871	Rattus norvegicus
77	16	1.1	5688	1 V19008	Rattus norvegicus
78	16	1.1	5910	1 V52298	Streptococcus pneu
79	16	1.1	6030	1 V19007	Rattus norvegicus
80	16	1.1	6039	1 V19870	Rattus norvegicus
81	16	1.1	7091	1 T47530	H6 promoted cyto
82	16	1.1	7696	1 X13474	Enterococcus faeca
83	16	1.1	9848	1 X20503	Polynucleotide seq
84	16	1.1	10095	1 V58938	Mycobacterium tube
85	16	1.1	10670	1 T99449	Staphylococcus car
86	16	1.1	11266	1 X12991	Enterococcus faeca
87	16	1.1	140875	1 T80043	Insert from cosmid
88	16	1.1	44377	1 T78508	Platenolide syntha
89	16	1.1	44377	1 T80414	Platenolide syntha
90	16	1.1	53789	1 V21187	Amycolatopsis medi
91	16	1.1	110000	1 V30458_2	Continuation (3 of
92	16	1.1	110000	1 V21209_03	Continuation (4 of
93	16	1.1	110000	1 V30459_2	Continuation (3 of
94	15	1.0	333	1 T17150	Human guanlylate ki
95	15	1.0	33	1 X15370	Human genome fragm
96	15	1.0	146	1 Q76949	Human genome fragm
97	15	1.0	151	1 V13097	Xylanase activity
98	15	1.0	352	1 Q37079	T cell receptor al
99	15	1.0	352	1 Q37080	T cell receptor al
100	15	1.0	352	1 Q37081	T cell receptor al
101	15	1.0	352	1 Q37082	T cell receptor al
102	15	1.0	385	1 T35130	Constant expressio
103	15	1.0	400	1 V78089	Staphylococcus aur
104	15	1.0	405	1 X41252	Human secreted pro
105	15	1.0	419	1 X51878	Human secreted pro
106	15	1.0	432	1 V90309	EST clone DK150. N

c 107	15	1.0	437	065637	Rat retrotransposo	180	15	1.0	1867	1	T39605	Mouse cortactin ge
c 108	15	1.0	481	X51674	Human secreted pro	c 181	15	1.0	1893	1	T64960	Twik-1 potassiu
c 109	15	1.0	513	T67435	H. pylori protein	c 182	15	1.0	1903	1	V34198	Human secreted p
c 110	15	1.0	515	V23862	Plant pox enzyme D	c 183	15	1.0	1997	1	T84276	Human Flt4 recept
c 111	15	1.0	527	Q45778	Human placenta tra	c 184	15	1.0	1997	1	V52576	Human vascular en
c 112	15	1.0	527	T14826	Human placenta tr	c 185	15	1.0	2002	1	T12320	Anti-procathesin
c 113	15	1.0	550	V63932	Mycobacterium tube	c 186	15	1.0	2012	1	Q43249	Sequence of Act20
c 114	15	1.0	606	T17148	Human guanlylate k1	c 187	15	1.0	2012	1	Q71448	Human vascular ha
c 115	15	1.0	606	X15376	Human guanlylate k1	c 188	15	1.0	2031	1	T59929	Human vascular ha
c 116	15	1.0	624	V23316	Plant PAL enzyme D	c 189	15	1.0	2037	1	Q71298	Probe for human ha
c 117	15	1.0	625	X15927	Lilly IgG1 nucleotl	c 190	15	1.0	2104	1	V60021	Nucleic acid G5ccc
c 118	15	1.0	625	X15927	Leishmania donovan	c 191	15	1.0	2120	1	V59660	Human secreted p
c 119	15	1.0	676	V47556	Leishmania antigen	c 192	15	1.0	2168	1	X52271	Protein PRO330 cDN
c 120	15	1.0	684	V23865	Plant PAL enzyme D	c 193	15	1.0	2175	1	X24906	Mouse elongation f
c 121	15	1.0	705	X15870	cDNA encoding inte	c 194	15	1.0	2178	1	X24905	Human elongation f
c 122	15	1.0	736	V83172	Human zeyto7 cDNA.	c 195	15	1.0	2207	1	V00500	Human alpha-(2) su
c 123	15	1.0	737	X02169	Human transcriptio	c 196	15	1.0	2222	1	O68556	Murine pro-hormone
c 124	15	1.0	789	V71949	Cancer marker MUC-	c 197	15	1.0	2254	1	O66167	Seven transmembran
c 125	15	1.0	798	T85783	Human Mucin 8 cDNA	c 198	15	1.0	2254	1	V18343	Human V28 seven tr
c 126	15	1.0	804	T38408	Mitochondrial ATP	c 199	15	1.0	2264	1	V81267	Bcl-2 interaction
c 127	15	1.0	804	T38403	ATP-synthase F1-Ar	c 200	15	1.0	2284	1	T31461	GABA-A receptor al
c 128	15	1.0	834	Q14087	Lipase gene. Recom	c 201	15	1.0	2310	1	O33129	GABA-A receptor al
c 129	15	1.0	844	X40000	Prostate cancer as	c 202	15	1.0	2310	1	O69141	Human GABA recept
c 130	15	1.0	891	V37169	DNA sequence used	c 203	15	1.0	2345	1	Q10868	KGF receptor gene.
c 131	15	1.0	945	X19530	Lilly Lgc1 promoter	c 204	15	1.0	2345	1	Q14049	Human ect gene. Ne
c 132	15	1.0	946	X40002	Prostate cancer as	c 205	15	1.0	2348	1	X05231	HIV-1 genomic DNA
c 133	15	1.0	960	T88154	S. marcescens Sep-h	c 206	15	1.0	2374	1	V18253	Bovine growth-indu
c 134	15	1.0	960	T88155	S. marcescens Sep-h	c 207	15	1.0	2556	1	V02872	Human htaflil05 cD
c 135	15	1.0	987	X04331	Human secreted pro	c 208	15	1.0	2675	1	T63340	Human host cell pr
c 136	15	1.0	990	T94210	Thermotoga maritim	c 209	15	1.0	2729	1	O51296	eIF-2 alpha kinase
c 137	15	1.0	1065	X27939	Human I(3)mbt prot	c 210	15	1.0	2729	1	Q44511	Haem-regulated euk
c 138	15	1.0	1089	X00638	Human secreted pro	c 211	15	1.0	2760	1	X13350	Enterococcus faeca
c 139	15	1.0	1127	X39991	Prostate cancer as	c 212	15	1.0	2770	1	T75035	Chick fringe B (luu
c 140	15	1.0	1139	T98609	DNA encoding a S.	c 213	15	1.0	2837	1	V10266	Rat GABA-BRB rece
c 141	15	1.0	1161	V18352	Human V28 seven tr	c 214	15	1.0	2911	1	V02311	C16 gene for promo
c 142	15	1.0	1162	O66170	Seven transmembran	c 215	15	1.0	2982	1	N91185	BamH1 insert from
c 143	15	1.0	1181	N90363	Limulus polyphemus	c 216	15	1.0	2991	1	V52588	Vascular endothell
c 144	15	1.0	1260	V41730	Codon-optimised RA	c 217	15	1.0	3012	1	X05825	Rat pheromone recee
c 145	15	1.0	1260	X05800	Mouse MAPK kinase	c 218	15	1.0	3065	1	V02312	C16n gene for prom
c 146	15	1.0	1289	Q09001	Bovine herpes viru	c 219	15	1.0	3070	1	O62694	Human G1UR3B rece
c 147	15	1.0	1302	T39789	Mouse SH3P7 gene.	c 220	15	1.0	3125	1	N70455	5'-3' sequence of
c 148	15	1.0	1323	V72290	S. mutans laab DNA	c 221	15	1.0	3286	1	X52267	Protein PRO332 cDN
c 149	15	1.0	1331	Q24681	H23-ETA-S antigen	c 222	15	1.0	3300	1	T62589	Thermotoga neophil
c 150	15	1.0	1356	T67851	H. pylori protein	c 223	15	1.0	3359	1	X05824	Rat pheromone recee
c 151	15	1.0	1370	T13022	Tyrosine phenillya	c 224	15	1.0	3394	1	X27920	Human I(3)mbt prot
c 152	15	1.0	1371	T19536	Xenopus DNA replic	c 225	15	1.0	3435	1	T35869	Human DNA polymera
c 153	15	1.0	1396	N80129	Limb gene-contg s	c 226	15	1.0	3442	1	N60705	Plasmid pIN603 enc
c 154	15	1.0	1488	Q38239	HCV CRS-NS5G recom	c 227	15	1.0	3512	1	X27921	Human I(3)mbt prot
c 155	15	1.0	1488	Q38234	HCV CRS-NS5G recom	c 228	15	1.0	3727	1	V11126	Rat SH2 binding pr
c 156	15	1.0	1488	Q38269	HCV CRS-NS5G fusio	c 229	15	1.0	3762	1	T73334	Human neuron-deriv
c 157	15	1.0	1491	Q21681	PHCV-51 - recombin	c 230	15	1.0	3976	1	V19872	Human sapiens cdo t
c 158	15	1.0	1503	X34119	Mycobacterium spec	c 231	15	1.0	3997	1	V24310	Human sapiens Class
c 159	15	1.0	1507	Q14003	Human cannabinoid	c 232	15	1.0	4054	1	O11888	Collagenase gene.
c 160	15	1.0	1514	V65632	Human heart O-fuco	c 233	15	1.0	4131	1	T89290	Dogfish shark kidn
c 161	15	1.0	1518	X05979	Mouse MAPK kinase	c 234	15	1.0	4190	1	O53530	Sequence of the un
c 162	15	1.0	1518	X05982	Mouse mutant MAPK	c 235	15	1.0	4212	1	V01098	Human propyl-4-hydr
c 163	15	1.0	1573	Q68222	Neuroendocrine tum	c 236	15	1.0	4310	1	V34856	Drosophila melano
c 164	15	1.0	1582	T84278	Human Flt4 recepto	c 237	15	1.0	4376	1	V10264	Rat GABA-BRIA recee
c 165	15	1.0	1583	V58758	Human secreted pro	c 238	15	1.0	4411	1	X28088	Plasmid pACYC. tac.
c 166	15	1.0	1587	X34120	Mycobacterium spec	c 239	15	1.0	4628	1	X13368	Enterococcus faeca
c 167	15	1.0	1598	X07064	Mouse mtogen acti	c 240	15	1.0	5009	1	V65634	First EcorI nucleo
c 168	15	1.0	1643	X07062	Mouse mtogen acti	c 241	15	1.0	5074	1	V02301	Rat tumour suppress
c 169	15	1.0	1683	T38409	Mitochondrial ATP	c 242	15	1.0	5175	1	X40200	Restin encoding ge
c 170	15	1.0	1683	T38404	ATP-synthase F1-Ar	c 243	15	1.0	5200	1	V02302	Rat tumour suppress
c 171	15	1.0	1704	T11413	Sequence encoding	c 244	15	1.0	5277	1	X13199	Enterococcus faeca
c 172	15	1.0	1709	Q29277	Encodes secreted f	c 245	15	1.0	5399	1	Q25706	pDEL08. Transforml
c 173	15	1.0	1738	V09299	Nucleotide sequenc	c 246	15	1.0	5454	1	X13343	Enterococcus faeca
c 174	15	1.0	1757	Q51034	Enhancer for the h	c 247	15	1.0	5615	1	V38689	Human sapiens SOCS1
c 175	15	1.0	1771	X40166	WO9904265 Seq ID N	c 248	15	1.0	5637	1	V57163	Partial human Notc
c 176	15	1.0	1795	T96035	Human kidney injur	c 249	15	1.0	5838	1	V28398	Plasmid pET14 nuc
c 177	15	1.0	1808	V02313	C16n gene for prom	c 250	15	1.0	5858	1	O10378	Encodes human 160k
c 178	15	1.0	1842	V34236	Human secreted pro	c 251	15	1.0	5889	1	V31499	Murine multidrug r
c 179	15	1.0	1844	V64372	GABA-gated chlorid	c 252	15	1.0	5889	1	X21978	Mouse multidrug re

c 253	15	1.0	5889	1	X19819	326	14	0.9	63	1	N81709	Insert from pSS100
c 254	15	1.0	6093	1	X25775	327	14	0.9	66	1	O58636	Encodes Bacillus m
c 255	15	1.0	6363	1	X28091	c 328	14	0.9	69	1	V64760	HIV anti-viral oli
c 256	15	1.0	6555	1	O33874	c 329	14	0.9	69	1	V64761	HIV anti-viral oli
c 257	15	1.0	6596	1	V57377	c 330	14	0.9	76	1	O38357	Sequence of RNA po
c 258	15	1.0	7399	1	O33479	c 331	14	0.9	76	1	O48832	Stem-loop confg. R
c 259	15	1.0	7399	1	V08704	c 332	14	0.9	76	1	V36649	Nucleotide sequenc
c 260	15	1.0	7650	1	V82019	c 333	14	0.9	76	1	V81183	RNA portion of a S
c 261	15	1.0	8091	1	V57001	c 334	14	0.9	78	1	V64768	HIV anti-viral oli
c 262	15	1.0	8752	1	V31189	c 335	14	0.9	98	1	V76384	Friedreich's ataxi
c 263	15	1.0	8906	1	V32979	c 336	14	0.9	108	1	V77627	Staphylococcus aur
c 264	15	1.0	9229	1	N80437	c 337	14	0.9	111	1	V15758	Ad 5 packaging sig
c 265	15	1.0	9229	1	V63467	c 338	14	0.9	138	1	T07037	Immunogen DNA from
c 266	15	1.0	9424	1	X23525	c 339	14	0.9	139	1	X27017	Differentially exp
c 267	15	1.0	9709	1	O22488	c 340	14	0.9	140	1	X54722	IL-8 receptor alph
c 268	15	1.0	9709	1	O96140	c 341	14	0.9	141	1	O33790	Downstream sequenc
c 269	15	1.0	9709	1	V81871	c 342	14	0.9	148	1	V15759	Ad 5 packaging sig
c 270	15	1.0	9738	1	X04765	c 343	14	0.9	152	1	O24206	Transglutaminase p
c 271	15	1.0	9739	1	O14751	c 344	14	0.9	154	1	O24885	PCR-generated prob
c 272	15	1.0	9739	1	V58549	c 345	14	0.9	156	1	V86799	EST clone BG309. N
c 273	15	1.0	9745	1	X04766	c 346	14	0.9	158	1	T01052	Mouse B7-1 gene pr
c 274	15	1.0	9746	1	O14752	c 347	14	0.9	159	1	X51473	Human secreted pro
c 275	15	1.0	9746	1	V58550	c 348	14	0.9	184	1	V03552	DNA sequence that
c 276	15	1.0	9822	1	N91774	c 349	14	0.9	189	1	O75182	AF-4 gene fragment
c 277	15	1.0	9960	1	V58939	c 350	14	0.9	218	1	V81417	Human prenylcyste
c 278	15	1.0	10091	1	V69642	c 351	14	0.9	241	1	T89402	Methods for diagno
c 279	15	1.0	10093	1	V58938	c 352	14	0.9	245	1	T41790	HIV-DRB1*08 allel
c 280	15	1.0	11284	1	V65633	c 353	14	0.9	245	1	T41794	HIV-DRB1*08 allel
c 281	15	1.0	11820	1	V81830	c 354	14	0.9	261	1	T28173	Sequence-related
c 282	15	1.0	12494	1	V83206	c 355	14	0.9	261	1	T21148	Human gene signatu
c 283	15	1.0	12588	1	O63293	c 356	14	0.9	267	1	O59419	Human brain expres
c 284	15	1.0	13256	1	X20528	c 357	14	0.9	269	1	T41792	HIV-DRB1*08 allel
c 285	15	1.0	14985	1	T43574	c 358	14	0.9	269	1	T41793	HIV-DRB1*08 allel
c 286	15	1.0	14985	1	V15693	c 359	14	0.9	269	1	T41795	HIV-DRB1*08 allel
c 287	15	1.0	15079	1	O91580	c 360	14	0.9	269	1	T41791	HIV-DRB1*08 allel
c 288	15	1.0	15464	1	O11415	c 361	14	0.9	269	1	T63554	Xylanase gene frag
c 289	15	1.0	15581	1	V14355	c 362	14	0.9	276	1	V66572	Mouse novel secret
c 290	15	1.0	15894	1	V18266	c 363	14	0.9	282	1	O13216	HIV-DNA6 coding s
c 291	15	1.0	20986	1	V52187	c 364	14	0.9	282	1	X40463	Human secreted pro
c 292	15	1.0	23677	1	X16323	c 365	14	0.9	284	1	O77212	Human genome fragm
c 293	15	1.0	24417	1	T97221	c 366	14	0.9	284	1	X51997	Human secreted pro
c 294	15	1.0	24601	1	X13160	c 367	14	0.9	285	1	V54994	Minimal CMV promot
c 295	15	1.0	28804	1	T37329	c 368	14	0.9	286	1	V88319	EST clone GD358. N
c 296	15	1.0	28804	1	T92474	c 369	14	0.9	288	1	T63553	Xylanase gene frag
c 297	15	1.0	28804	1	V81474	c 370	14	0.9	289	1	V88814	EST clone HO1216.
c 298	15	1.0	28804	1	V99812	c 371	14	0.9	291	1	T35327	Pancreas expressed
c 299	15	1.0	30001	1	T61016	c 372	14	0.9	294	1	T58777	Human eotaxin cDNA
c 300	15	1.0	30001	1	X05110	c 373	14	0.9	296	1	V89317	EST clone C1437. N
c 301	15	1.0	35100	1	V20441	c 374	14	0.9	298	1	O61372	Human brain expres
c 302	15	1.0	110000	1	T58840_1	c 375	14	0.9	303	1	V49597	Human Osteocarcin
c 303	15	1.0	133894	1	T13635	c 376	14	0.9	308	1	V21411	Lambda phage heter
c 304	15	1.0	235033	1	V57926	c 377	14	0.9	315	1	V88416	EST clone EK581. N
c 305	15	1.0	237326	1	V57903	c 378	14	0.9	318	1	O60881	Human brain expres
c 306	14	0.9	18	1	T01361	c 379	14	0.9	324	1	N61362	Sequence of alpha-
c 307	14	0.9	18	1	V12406	c 380	14	0.9	325	1	X26289	5' end of human Rp
c 308	14	0.9	19	1	V61129	c 381	14	0.9	327	1	V67698	EST clone EN423. N
c 309	14	0.9	20	1	O26368	c 382	14	0.9	333	1	X41120	Human secreted pro
c 310	14	0.9	20	1	O76007	c 383	14	0.9	336	1	O43541	AF-4/ALF-1 exon 9
c 311	14	0.9	20	1	V08298	c 384	14	0.9	336	1	O75184	ALF-1/AF-4 chimeri
c 312	14	0.9	21	1	T30450	c 385	14	0.9	338	1	T26484	Human gene signatu
c 313	14	0.9	23	1	X00022	c 386	14	0.9	347	1	X41118	Human secreted pro
c 314	14	0.9	23	1	X00010	c 387	14	0.9	387	1	T91527	Mycobacterium tube
c 315	14	0.9	26	1	T58780	c 388	14	0.9	387	1	T91461	Mycobacterium tube
c 316	14	0.9	27	1	T34089	c 389	14	0.9	387	1	V44401	M. tuberculosis im
c 317	14	0.9	27	1	V71708	c 390	14	0.9	389	1	V78553	Staphylococcus aur
c 318	14	0.9	29	1	T39359	c 391	14	0.9	389	1	O61266	Human brain expres
c 319	14	0.9	30	1	T86252	c 392	14	0.9	394	1	X41296	Human secreted pro
c 320	14	0.9	33	1	O45859	c 393	14	0.9	395	1	O59672	Human brain expres
c 321	14	0.9	36	1	O52042	c 394	14	0.9	399	1	V05729	Nucleotide sequenc
c 322	14	0.9	40	1	T67172	c 395	14	0.9	400	1	V78222	Staphylococcus aur
c 323	14	0.9	40	1	V08300	c 396	14	0.9	400	1	X54718	IL-8 receptor alph
c 324	14	0.9	41	1	V50709	c 397	14	0.9	419	1	T22604	Human gene signatu
c 325	14	0.9	57	1	T60640	c 398	14	0.9	419	1		

C 399	14	0.9	419	1	V32960	Human ZPPAR6 nucle	C 472	14	0.9	631	1	N70531	Replicon and contr
C 400	14	0.9	425	1	V52520	Streptococcus pneu	C 473	14	0.9	632	1	V16847	Tobacco mosaic vir
C 401	14	0.9	430	1	X41276	Human secreted pro	C 474	14	0.9	633	1	V88133	EST clone FY455, N
C 402	14	0.9	432	1	T91525	Mycobacterium tube	C 475	14	0.9	639	1	Q20764	Human ITRNalpha 2C
C 403	14	0.9	432	1	T91459	Mycobacterium tube	C 476	14	0.9	639	1	Q20764	Human ITRNalpha 2C
C 404	14	0.9	432	1	V44389	Mycobacterium tube	C 477	14	0.9	642	1	T89388	PAD19B-1FN HindIII
C 405	14	0.9	432	1	V64508	M. tuberculosis im	C 478	14	0.9	647	1	Q47153	Methods for diagno
C 406	14	0.9	450	1	Q76266	PtCMV*-1 promoter.	C 479	14	0.9	651	1	Q48029	Natural human inte
C 407	14	0.9	450	1	Q76267	PtCMV*-2 promoter.	C 480	14	0.9	654	1	Q10269	Potato Leaf Roll V
C 408	14	0.9	450	1	T06869	PtCMV*-1 telo cons	C 481	14	0.9	660	1	N90371	Human BMP. Purifie
C 409	14	0.9	450	1	T06870	PtCMV*-2 telo cons	C 482	14	0.9	660	1	N81708	DNA encoding Bacil
C 410	14	0.9	450	1	T11355	Minimal CMV promot	C 483	14	0.9	662	1	V74698	Multiple promoter
C 411	14	0.9	450	1	T11356	Minimal CMV promot	C 484	14	0.9	662	1	X20061	Streptococcus aur
C 412	14	0.9	450	1	V60080	Cytomegalovirus m1	C 485	14	0.9	665	1	V74935	Enterococcus faeca
C 413	14	0.9	450	1	V60081	Cytomegalovirus m1	C 486	14	0.9	674	1	O53499	Staphylococcus aur
C 414	14	0.9	450	1	X01365	PtCMV*-1 sequence.	C 487	14	0.9	675	1	O53500	Sequence encoding
C 415	14	0.9	450	1	X01367	PtCMV*-2 sequence.	C 488	14	0.9	690	1	V52510	Sequence encoding
C 416	14	0.9	450	1	X27901	PtCMV*-1 promoter.	C 489	14	0.9	690	1	V49596	Streptococcus pneu
C 417	14	0.9	450	1	X27903	PtCMV*-2 promoter.	C 490	14	0.9	712	1	V47588	Human Osteocarcin
C 419	14	0.9	462	1	V33342	Synthetic human in	C 491	14	0.9	720	1	O11142	Alkaline phosphata
C 420	14	0.9	481	1	V17755	Inducible promoter	C 492	14	0.9	741	1	Q45080	Human glucokinase
C 421	14	0.9	481	1	V18691	Cytomegalovirus im	C 493	14	0.9	744	1	N10011	Sequence of the H1
C 422	14	0.9	486	1	Q14344	Avr9 avirulence ge	C 494	14	0.9	750	1	V52503	Streptococcus pneu
C 423	14	0.9	488	1	N50236	Sequence encoding	C 495	14	0.9	750	1	X32470	C. amalonaticus ma
C 424	14	0.9	498	1	N60747	Sequence encoding	C 496	14	0.9	757	1	T01587	Her-2/neu (ERBB2/C
C 425	14	0.9	500	1	N70528	EST clone AB45. Ne	C 497	14	0.9	758	1	V18170	Gulf War veteran s
C 426	14	0.9	502	1	X07001	Sequence encoding	C 498	14	0.9	759	1	V18172	Gulf War veteran s
C 427	14	0.9	503	1	N40013	Xenopus Frazzled p	C 499	14	0.9	760	1	V18171	Gulf War veteran s
C 428	14	0.9	504	1	Q47152	DNA encoding recom	C 500	14	0.9	765	1	T17258	Oil body protein g
C 429	14	0.9	504	1	V52522	Recombinant human	C 501	14	0.9	765	1	X20060	Enterococcus faeca
C 430	14	0.9	505	1	N80047	Streptococcus pneu	C 502	14	0.9	771	1	T14068	Human cytoegalov1
C 431	14	0.9	506	1	N60654	Sequence encoding	C 503	14	0.9	771	1	T14064	Human cytoegalov1
C 432	14	0.9	506	1	N30155	Encodes human IFN-	C 504	14	0.9	771	1	T14065	Human cytoegalov1
C 433	14	0.9	519	1	Q25000	Sequence encoding	C 505	14	0.9	774	1	T28667	Human cytoegalov1
C 434	14	0.9	519	1	Q24989	Human interferon a	C 506	14	0.9	780	1	N70958	Ripening banana pu
C 435	14	0.9	519	1	Q40478	Human interferon a	C 507	14	0.9	793	1	N20008	Sequence of the 5'
C 436	14	0.9	520	1	T11359	MUREG-2 coding reg	C 508	14	0.9	793	1	V65198	Hybrid human Leuko
C 437	14	0.9	520	1	T45722	Bi-directional prom	C 509	14	0.9	802	1	N60312	Human mitochondria
C 438	14	0.9	520	1	V60079	Bi-directional tet	C 510	14	0.9	803	1	V636303	Bequine IFN-omega-2
C 439	14	0.9	540	1	O11141	Bi-directional prom	C 511	14	0.9	806	1	N60113	Human guanine nucl
C 440	14	0.9	545	1	V88972	Interferon alpha a	C 512	14	0.9	806	1	N40113	Sequence of the co
C 441	14	0.9	548	1	T75095	EST clone J101. N	C 513	14	0.9	807	1	V52505	Sequence of the cd
C 442	14	0.9	561	1	T89462	Sinapis alpha flowe	C 514	14	0.9	808	1	N30106	Streptococcus pneu
C 443	14	0.9	561	1	T45894	Flatfish insulin-1	C 515	14	0.9	828	1	T74477	Sequence on recomb
C 444	14	0.9	568	1	N80043	R. capsulatus acet	C 516	14	0.9	831	1	X39996	Open reading frame
C 445	14	0.9	568	1	N60652	Sequence encoding	C 517	14	0.9	840	1	T37431	Prostate cancer as
C 446	14	0.9	568	1	N30151	HindIII-EcoRI frag	C 518	14	0.9	843	1	X07061	Bali mottification
C 447	14	0.9	569	1	T11358	Sequence encoding	C 519	14	0.9	847	1	V23492	Human mitogen acti
C 448	14	0.9	569	1	T11358	Bi-directional prom	C 520	14	0.9	859	1	T62944	Pseudomonas Orf2 s
C 449	14	0.9	569	1	T45721	Bi-directional prom	C 521	14	0.9	863	1	N20093	Human eosinocyte c
C 450	14	0.9	569	1	T45721	Bi-directional tet	C 522	14	0.9	869	1	V52502	Sequence of leukoc
C 451	14	0.9	569	1	V60078	Bi-directional tet	C 523	14	0.9	874	1	N10010	Streptococcus pneu
C 452	14	0.9	569	1	V60078	Bi-directional prom	C 524	14	0.9	879	1	N40110	Sequence of the H1
C 453	14	0.9	573	1	Q04744	Sequence encoding	C 525	14	0.9	879	1	Q79523	Sequence of the cd
C 454	14	0.9	573	1	T16625	Hepatitis C virus	C 526	14	0.9	880	1	N30103	PhoA promoter-STII
C 455	14	0.9	583	1	T16630	Hepatitis C virus	C 527	14	0.9	895	1	O62653	Sequence on recomb
C 456	14	0.9	586	1	Q23083	Antigen tc-23g gen	C 528	14	0.9	897	1	V37165	Tobacco-mosaic vir
C 457	14	0.9	586	1	V52518	Streptococcus pneu	C 529	14	0.9	903	1	Q23361	ADRI gene from Can
C 458	14	0.9	586	1	X22088	Streptococcus pneu	C 530	14	0.9	913	1	X20348	ADRI gene from Can
C 459	14	0.9	588	1	N50272	Potato G023586 RAPD	C 531	14	0.9	915	1	T70161	Borrelia burgdorfe
C 460	14	0.9	594	1	N50273	DNA encoding inter	C 532	14	0.9	919	1	N90695	Scal methylase gen
C 461	14	0.9	600	1	N93069	DNA encoding inter	C 533	14	0.9	919	1	N81527	DNA sequence of po
C 462	14	0.9	604	1	T75096	Fragment of pCT-el	C 534	14	0.9	925	1	V99229	Sequence of a port
C 463	14	0.9	604	1	X03393	Arabidopsis flower	C 535	14	0.9	927	1	V74985	DNA encoding an ac
C 464	14	0.9	611	1	V52735	Coding sequence fo	C 536	14	0.9	930	1	N60156	Staphylococcus aur
C 465	14	0.9	616	1	Q53551	Human hepatocyte n	C 537	14	0.9	930	1	Q43524	Enhancer for enkar
C 466	14	0.9	616	1	T77192	HCNV (Towne) major	C 538	14	0.9	930	1	Q43525	Sequence of HCNV/A
C 467	14	0.9	617	1	V42947	HCNV immediate ear	C 539	14	0.9	930	1	O53550	Sequence of HCNV/T
C 468	14	0.9	620	1	V52507	Streptococcus pneu	C 540	14	0.9	930	1	T77193	HCNV (AD169) major
C 469	14	0.9	627	1	Q37629	Streptococcus pneu	C 541	14	0.9	941	1	N30062	HCNV immediate ear
C 470	14	0.9	627	1	Q37630	Native PLRV coat p	C 542	14	0.9	941	1	V52501	Sequence of a modl
C 471	14	0.9	627	1	Q53460	Modified PLRV coat	C 543	14	0.9	948	1	Q80574	Streptococcus pneu
						PLRV Integument pr	C 544	14	0.9	948	1	T33294	HIV-1 pol gene. Me

Thu May 11 11:19:43 2000

us-09-080-127-1.oligo.rng

Page 5

545	14	0.9	954	1	X20346	Borrelia burgdorferi
c 546	14	0.9	958	1	N20090	Sequence of leukoc
c 547	14	0.9	958	1	N20005	Hybrid human leuko
c 548	14	0.9	964	1	Q25755	Non-A, Non-B Hepat
c 549	14	0.9	981	1	N30045	Sequence encoding
c 550	14	0.9	981	1	V68166	Bovine Interferon-
c 551	14	0.9	981	1	V08183	Bovine Interferon-
c 552	14	0.9	990	1	Q12852	Hybrid Interferon
c 553	14	0.9	990	1	V68080	DNA encoding an ox
554	14	0.9	991	1	T85884	Malassezia fungus
c 555	14	0.9	998	1	Q04548	Residual protease-
556	14	0.9	998	1	Q87815	xpa1/xpa2 coding
c 557	14	0.9	998	1	T59387	RP-I gene (mpri) en
c 558	14	0.9	998	1	T63356	Bacillus subtilis
c 559	14	0.9	999	1	T90124	Human CCR4L (TRAF-
560	14	0.9	1000	1	Q04473	Potato Leaf roll v
c 561	14	0.9	1014	1	Q68438	Pseudomonas glutam
c 562	14	0.9	1035	1	T49316	Hypersensitive res
c 563	14	0.9	1035	1	V36430	Hypersensitive res
c 564	14	0.9	1035	1	V39975	Pseudomonas solana
c 565	14	0.9	1035	1	V54609	Pseudomonas solana
c 566	14	0.9	1035	1	V83991	DNA encoding a hyp
c 567	14	0.9	1039	1	V68168	Porcine Interferon
c 568	14	0.9	1051	1	V08185	Porcine Interferon
569	14	0.9	1051	1	V62994	Human HOEC11 cDNA
c 570	14	0.9	1054	1	T61257	NA17-A cDNA encodi
c 571	14	0.9	1068	1	V52506	Streptococcus pneu
572	14	0.9	1070	1	X29176	M. tuberculosis re
573	14	0.9	1078	1	T01287	Rh1pcephalus appe
574	14	0.9	1087	1	T98694	DNA encoding a S.
c 575	14	0.9	1090	1	X37521	Human secreted pro
576	14	0.9	1090	1	V52664	Human native hepat
c 577	14	0.9	1109	1	N20026	Human leukocyte in
c 578	14	0.9	1122	1	N92362	DNA sequence encod
579	14	0.9	1140	1	V41733	Codon-optimised RA
c 580	14	0.9	1140	1	X06999	Xenopus Frazzled P
c 581	14	0.9	1143	1	V29233	Nucleotide sequenc
c 582	14	0.9	1146	1	Q86990	Creatine-kinase su
583	14	0.9	1158	1	V42043	Human tailless nuc
c 584	14	0.9	1161	1	Q99489	3 hydroxy-beta-ion
c 585	14	0.9	1161	1	V84076	Carotenoid biosynt
c 586	14	0.9	1161	1	X19116	Agrobacterium aura
c 587	14	0.9	1162	1	N30044	Sequence encoding
c 588	14	0.9	1162	1	V68165	Bovine Interferon-
c 589	14	0.9	1162	1	V08182	Recombinant high a
c 590	14	0.9	1176	1	Q99950	Sequence of the Hc
c 591	14	0.9	1179	1	N10013	RBEL transcritpion
c 592	14	0.9	1181	1	V12742	S-adenosylmethioni
c 593	14	0.9	1182	1	T99141	S-adenosylmethioni
c 594	14	0.9	1182	1	T99143	S-adenosylmethioni
595	14	0.9	1183	1	Q25786	SHPI (serotonin) r
596	14	0.9	1212	1	N81271	Sequence encoding
c 597	14	0.9	1212	1	N60274	E. coli L-carnatine
c 598	14	0.9	1215	1	Q98553	DNA flanking the T
599	14	0.9	1216	1	T95021	CytJl ORF1 coding
c 600	14	0.9	1225	1	Q73386	American elm chlti
601	14	0.9	1225	1	T33325	Bacillus cereus st
c 602	14	0.9	1229	1	T14429	Nucleotide sequenc
c 603	14	0.9	1236	1	V48114	Nucleotide sequenc
c 604	14	0.9	1239	1	V39014	Human utrophin gen
c 605	14	0.9	1246	1	T45971	Murine 4-1BB-L pol
c 606	14	0.9	1254	1	Q75422	Murine 4-1BB-L pol
c 607	14	0.9	1254	1	T91024	Fragment of IBM ligan
c 608	14	0.9	1256	1	T91024	Human Growth hormo
609	14	0.9	1257	1	Q45310	S. cerevisiae frag
610	14	0.9	1263	1	T04490	S. cerevisiae frag
611	14	0.9	1263	1	T04491	Human growth hormo
612	14	0.9	1269	1	T61022	Arabidopsis JAG18
613	14	0.9	1271	1	T93020	Human Growth hormo
614	14	0.9	1272	1	Q45311	33 pGRF receptor g
615	14	0.9	1272	1	Q66344	Human acromegalic
616	14	0.9	1272	1	T47109	HAP 7.3 DNA encodi
617	14	0.9	1272	1	T77198	
618	14	0.9	1272	1	V81456	Human clone HAP7.3
619	14	0.9	1278	1	T63617	Rice tapetum speci
620	14	0.9	1278	1	X29158	Regulatory polynuc
621	14	0.9	1289	1	T97166	Mouse HPR1 gene. M
622	14	0.9	1290	1	Q66343	pGRF receptor gene
c 623	14	0.9	1302	1	T69920	DNA encoding inter
c 624	14	0.9	1308	1	X05978	Human MAPK kinase
c 625	14	0.9	1320	1	T72050	Sequence around ma
c 626	14	0.9	1320	1	T43930	Sequence around ma
c 627	14	0.9	1320	1	V03316	Thermococcus celer
628	14	0.9	1320	1	V20948	Coffee-fruit speci
c 629	14	0.9	1326	1	T61729	Streptococcus pneu
c 630	14	0.9	1347	1	T41700	Brugia pahangi bet
c 631	14	0.9	1356	1	Q66342	pGRF receptor gene
632	14	0.9	1360	1	V27342	Streptococcus pneu
c 633	14	0.9	1366	1	Q36768	Sequence of glycop
c 634	14	0.9	1386	1	Q36768	Sequence of glycop
c 635	14	0.9	1386	1	Q97670	Infectious bovine
c 636	14	0.9	1393	1	Q97670	Infectious bovine
637	14	0.9	1393	1	V21410	Lambda phage heter
c 638	14	0.9	1396	1	V18393	Human immunodefici
639	14	0.9	1400	1	Q62460	Bacillus subtilis
c 640	14	0.9	1400	1	T45418	B. subtilis entero
c 641	14	0.9	1425	1	X30785	Streptococcus pneu
c 642	14	0.9	1436	1	V52498	Streptococcus pneu
643	14	0.9	1437	1	Q36687	Sequence encoding
644	14	0.9	1437	1	Q48234	Cytochrome C gene.
645	14	0.9	1438	1	N91275	DNA encoding Asper
646	14	0.9	1441	1	V52687	Human native hepat
c 647	14	0.9	1448	1	T69574	Mouse tumour susce
648	14	0.9	1453	1	T61724	Streptococcus pneu
649	14	0.9	1458	1	V36418	Streptococcus pneu
650	14	0.9	1470	1	Q10869	Keratan sulphate 6
651	14	0.9	1472	1	Q98497	Rabbit 50 kDa dyst
652	14	0.9	1472	1	T97323	Rabbit (50 kDa) dy
c 653	14	0.9	1479	1	N60728	Sequence encoding
c 654	14	0.9	1485	1	X07184	Soybean S-adenosyl
c 655	14	0.9	1489	1	Q67073	DNA coding serine
c 656	14	0.9	1490	1	T89357	Methods for diagno
c 657	14	0.9	1494	1	V12204	Human retinoid met
c 658	14	0.9	1494	1	Q109247	Human cytochrome P
659	14	0.9	1495	1	V31965	B. campestris delt
660	14	0.9	1495	1	V15253	Brassica campestris
661	14	0.9	1497	1	Q03367	Phenol oxidase (PO
662	14	0.9	1497	1	Q03571	Sequence encoding
663	14	0.9	1497	1	Q03572	Sequence encoding
664	14	0.9	1497	1	V01915	Human GSP1-TK gen
c 665	14	0.9	1501	1	T89405	Methods for diagno
c 666	14	0.9	1505	1	T98625	DNA encoding a S.
c 667	14	0.9	1508	1	V57740	Partial human 7-tr
668	14	0.9	1526	1	V53522	DNA encoding a sta
669	14	0.9	1526	1	T45972	Mouse utrophin gen
670	14	0.9	1536	1	V42334	Trametes pubescens
c 671	14	0.9	1537	1	V24291	Pseudomonas aerugi
c 672	14	0.9	1541	1	T04133	Helicobacter pylori
c 673	14	0.9	1558	1	X22942	Human cathepsin Y
c 674	14	0.9	1566	1	T39768	Alpha-Interferon B
675	14	0.9	1568	1	Q51620	Human creatine kin
676	14	0.9	1568	1	Q58990	Mutant human creat
677	14	0.9	1575	1	Q12342	Actinomycete Phosp
c 678	14	0.9	1575	1	T94206	ARPIIIA clone 63CP
c 679	14	0.9	1577	1	T90169	Oil seed rape cyst
c 680	14	0.9	1581	1	Q79601	Human c-erbB-2 5'
c 681	14	0.9	1615	1	Q99201	Pseudomonas putida
c 682	14	0.9	1617	1	Q62312	Human somatoliberi
c 683	14	0.9	1619	1	Q04585	Fragmentary sequen
c 684	14	0.9	1624	1	Q85984	Oryza sativa pppK-
c 685	14	0.9	1633	1	T00819	Mouse type-II memb
c 686	14	0.9	1633	1	V32635	Murine OX40-L cDNA
c 687	14	0.9	1635	1	V27578	Human hepatocyte n
688	14	0.9	1642	1	Q14835	Human serotonin re
c 689	14	0.9	1649	1	X33111	Candida rugosa lip
c 690	14	0.9	1650	1	T10422	Candida cylindrace

691	14	0.9	1674	1	X2532	Human MMP-20 DNA.
692	14	0.9	1702	1	Q97236	Rhizobium dcla gen
c 693	14	0.9	1702	1	V52500	Streptococcus pneu
694	14	0.9	1704	1	O86456	Spleen necrosis vl
c 695	14	0.9	1713	1	N92533	CDNA Clone P107 en
c 696	14	0.9	1713	1	Q73002	Human leukocyte pl
697	14	0.9	1718	1	V15537	Human apolipoprote
698	14	0.9	1751	1	Q71193	Clavulanic acid re
699	14	0.9	1753	1	T51422	ESAT6 coding seque
c 700	14	0.9	1755	1	O49461	Neurokinin-3 recep
c 701	14	0.9	1759	1	T63616	Rice tapetum speci
702	14	0.9	1779	1	Q13322	PLRV Saci-EcorI fr
703	14	0.9	1785	1	X23404	Mouse scavenger re
c 704	14	0.9	1789	1	T34657	Nit protease of th
c 705	14	0.9	1791	1	X02189	Rat anti-oestrogen
706	14	0.9	1794	1	O83534	Sequence encoding
c 707	14	0.9	1800	1	X13723	Enterococcus faeca
c 708	14	0.9	1816	1	T68722	Human osteo ancli1
c 709	14	0.9	1825	1	Q67663	TMV replicon RNA.
c 710	14	0.9	1825	1	V54825	Replicon RNA produ
c 711	14	0.9	1831	1	X26880	DNA encoding a pro
c 712	14	0.9	1840	1	O81826	Alpha-N-acetylglu
c 713	14	0.9	1841	1	V45329	Human extracellular
714	14	0.9	1844	1	Q30175	63 kD CAM PDE CDNA
715	14	0.9	1844	1	O83969	Cyclic-GMP stimula
716	14	0.9	1844	1	T51105	CDNA for 63 kD cal
717	14	0.9	1844	1	T67200	Bovine brain 63 kD
718	14	0.9	1844	1	V36148	Bovine brain 63 kD
719	14	0.9	1844	1	V48196	Bovine brain Ca2+/
720	14	0.9	1844	1	V54753	63 kDa CAM-PDE DNA
c 721	14	0.9	1848	1	T66346	Chimeric gene cont
722	14	0.9	1851	1	V68992	DNA molecule encod
723	14	0.9	1851	1	V68993	DNA molecule encod
724	14	0.9	1851	1	V71305	DNA encoding ester
c 725	14	0.9	1863	1	N60399	Sequence of the pr
726	14	0.9	1864	1	T64931	Actinomadura flexu
727	14	0.9	1876	1	T87067	Chicken platelet-a
728	14	0.9	1876	1	V08553	Chicken PAR-AH cod
729	14	0.9	1876	1	X08479	Avian plasma plate
c 730	14	0.9	1882	1	T90522	Oxalyl-CoA decarbo
c 731	14	0.9	1882	1	T75517	CDNA encoding huma
c 732	14	0.9	1883	1	O99006	Interleukin 8 rece
c 733	14	0.9	1899	1	X05360	Human extracellular
c 734	14	0.9	1901	1	Q12853	BamHI-EcoRI fragme
c 735	14	0.9	1917	1	T09873	Human DNA-topoisom
c 736	14	0.9	1917	1	V15009	Human DNA-topoisom
c 737	14	0.9	1919	1	N90817	Rat NADH-cytochrom
c 738	14	0.9	1933	1	Q80520	Interleukin-8 rece
c 739	14	0.9	1951	1	T74893	Mouse neurogenic d
c 740	14	0.9	1951	1	V42934	CDNA encoding murI
c 741	14	0.9	1955	1	Q12854	BamHI-EcoRI fragme
c 742	14	0.9	1970	1	O13619	Nitrite hydratase
c 743	14	0.9	1980	1	O90197	VZV thymidine-kin
c 744	14	0.9	1980	1	O90199	VZV thymidine-kin
c 745	14	0.9	1980	1	O90201	VZV thymidine-kin
c 746	14	0.9	1980	1	T07314	Varicella zoster v
747	14	0.9	1983	1	T47497	Porcine growth hor
748	14	0.9	1990	1	T61725	Streptococcus pneu
749	14	0.9	2000	1	T58391	T. harzianum exoch
750	14	0.9	2000	1	V84063	CDNA encoding a he
751	14	0.9	2006	1	X20417	Human secreted pro
c 752	14	0.9	2021	1	Q11301	Leucocyte-plaslin
c 753	14	0.9	2036	1	V32538	Human granulate bl
c 754	14	0.9	2049	1	O03366	Phenol oxidase (PO
755	14	0.9	2049	1	O03573	Phenol oxidase gen
756	14	0.9	2057	1	V01916	Human GSP1-TK gen
c 757	14	0.9	2070	1	Q24875	Human amyloidln pr
758	14	0.9	2085	1	Q78131	Pneumococcal surfa
759	14	0.9	2085	1	T07178	Pneumococcal surfa
760	14	0.9	2085	1	T08979	DNA encoding pneum
761	14	0.9	2085	1	V39470	Streptococcus pneu
762	14	0.9	2085	1	V33264	Pneumococcal surfa
763	14	0.9	2085	1	V84069	DNA encoding a pne
764	14	0.9	2085	1	X02012	S. pneumoniae trun
765	14	0.9	2095	1	O57953	Human beta1,6-N-ac
766	14	0.9	2095	1	V12251	Novel protein from
767	14	0.9	2095	1	V15915	HIV-1 partial DNA
768	14	0.9	2095	1	X19801	HIV-1 DNA sequence
c 769	14	0.9	2107	1	X33944	Human HCMV Inducib
c 770	14	0.9	2133	1	N91042	Promoter-enhancer
c 771	14	0.9	2139	1	T41704	Lymphocyte specific
c 772	14	0.9	2158	1	O25156	Alpha-GalNac from
c 773	14	0.9	2179	1	Q70008	Fibronectin. New h
c 774	14	0.9	2184	1	V68997	DNA molecule encod
c 775	14	0.9	2190	1	X06998	Xenopus Frazzled p
c 776	14	0.9	2196	1	Q74211	CMVlnfa promoter a
c 777	14	0.9	2196	1	T01112	CMVlnfa-GBH termin
c 778	14	0.9	2214	1	O13580	A. allicoccigenes me
c 779	14	0.9	2214	1	Q20383	ADH complex struct
c 780	14	0.9	2225	1	T45101	Cytomegalovirus po
781	14	0.9	2230	1	V99806	Murine lysophospha
c 782	14	0.9	2256	1	T32999	Chimeric thermost
c 783	14	0.9	2258	1	V37374	Streptococcus pneu
c 784	14	0.9	2265	1	X26288	Human RPMI 8392 ce
785	14	0.9	2276	1	N93080	Formate dehydroge
786	14	0.9	2288	1	V27577	Human hepatocyte n
787	14	0.9	2289	1	V27579	Human hepatocyte n
788	14	0.9	2299	1	O14176	Clone pXR2C8 encod
789	14	0.9	2304	1	O55374	pXR2C8 DNA. Multim
790	14	0.9	2304	1	T76787	Insect XRCc recept
791	14	0.9	2304	1	T89959	D. melanogaster XR
792	14	0.9	2311	1	X13205	Cephalosporin C ac
793	14	0.9	2325	1	Q12578	Human islet cell a
c 794	14	0.9	2338	1	T95227	Retinoid X recepto
c 795	14	0.9	2389	1	T31930	Cephalosporin amid
796	14	0.9	2391	1	T29215	Protein PRO294 cDN
c 797	14	0.9	2403	1	X52254	Mycobacterium spec
798	14	0.9	2409	1	X34178	DNA encoding malic
799	14	0.9	2430	1	O83161	CDNA encoding glyco
800	14	0.9	2432	1	T77045	Human endothelin c
c 801	14	0.9	2440	1	T31399	Human endothelin c
c 802	14	0.9	2440	1	T42151	Mycobacterium spec
c 803	14	0.9	2451	1	X34179	CDNA of cbpA, a ch
804	14	0.9	2462	1	V05102	Farnesyltransferas
c 805	14	0.9	2464	1	O44398	Rat farnesyl prote
c 806	14	0.9	2464	1	O94411	Rat farnesyl trans
c 807	14	0.9	2464	1	T38709	Human heat shock p
c 808	14	0.9	2465	1	T58086	Equine IFN-beta fr
809	14	0.9	2467	1	M60307	CDNA encoding huma
c 810	14	0.9	2482	1	X26693	Human pituitary do
c 811	14	0.9	2483	1	O04790	CDNA encoding huma
c 812	14	0.9	2490	1	V16208	Human complement c
c 813	14	0.9	2492	1	X24297	Protein PRO262 cDN
c 814	14	0.9	2505	1	X52246	Human GTP binding
815	14	0.9	2516	1	V52606	Putative human cad
c 816	14	0.9	2521	1	T61926	Human cadherin-12
c 817	14	0.9	2521	1	T85404	Human placental en
c 818	14	0.9	2533	1	Q93877	Encodes human reti
c 819	14	0.9	2536	1	Q11621	Streptococcus pneu
c 820	14	0.9	2540	1	O37687	Actinoplanes sp. a
c 821	14	0.9	2550	1	V52497	Maize Lc CDNA sequ
c 822	14	0.9	2582	1	T73117	High oligallatly G1
c 823	14	0.9	2605	1	O10790	Human retinoid met
c 824	14	0.9	2605	1	V04699	Canine beta-2 adre
c 825	14	0.9	2624	1	T44328	Human secreted pro
c 826	14	0.9	2677	1	V12214	Moraxella catarin
c 827	14	0.9	2679	1	V30468	Human ECE gene. Ne
c 828	14	0.9	2705	1	V58756	Human endothelin c
c 829	14	0.9	2718	1	V82021	Human eotaxin geno
c 830	14	0.9	2720	1	O90315	Gastrin-binding pr
c 831	14	0.9	2730	1	O93861	Mouse vav proto on
c 832	14	0.9	2744	1	T58778	Human striated mus
833	14	0.9	2744	1	O31769	PKG-green fluoresc
834	14	0.9	2793	1	Q26682	
835	14	0.9	2793	1	V48318	
c 836	14	0.9	2799	1	V71085	

C 837	14	0.9	2802	1	V71084	Green fluorescent
C 838	14	0.9	2812	1	O55148	Pseudomonas aerugi
C 839	14	0.9	2826	1	O51528	Wheat granule-boun
C 840	14	0.9	2836	1	T69800	Aureobasidin sensi
C 841	14	0.9	2856	1	V33679	Aspergillus nidula
C 842	14	0.9	2877	1	V33680	Aspergillus nidula
C 843	14	0.9	2947	1	V18821	Murine hyaluronan
C 844	14	0.9	2970	1	X20290	Borrelia burgdorfe
C 845	14	0.9	2987	1	O22937	Sequence encoding
C 846	14	0.9	3023	1	V38817	C3 binding protein
C 847	14	0.9	3035	1	X14059	H. pylori GHP0 325
C 848	14	0.9	3066	1	O48985	Rat glucanase-like
C 849	14	0.9	3089	1	X07348	Human GTPase regul
C 850	14	0.9	3090	1	V38720	Full length cassav
C 851	14	0.9	3097	1	V19609	Homo sapiens EPCR
C 852	14	0.9	3125	1	O73686	HIV virus-1(LAI) 9
C 853	14	0.9	3147	1	T86704	DNA encoding therm
C 854	14	0.9	3171	1	V99828	Human guanine nucl
C 855	14	0.9	3182	1	T37432	Plasmid pBSRI inse
C 856	14	0.9	3200	1	MA0243	DNA encoding pre H
C 857	14	0.9	3223	1	V31353	Trichoderma reesei
C 858	14	0.9	3242	1	N90819	Rat NADH-cytochrom
C 859	14	0.9	3260	1	V58387	Hepatitase codin
C 860	14	0.9	3264	1	V26606	Plasmid pUC18FLA23
C 861	14	0.9	3402	1	T27310	Trypanosoma cruzi
C 862	14	0.9	3438	1	T95247	M. catarrhalis 422
C 863	14	0.9	3530	1	T03438	Human matrix metal
C 864	14	0.9	3547	1	T01115	Human matrix metal
C 865	14	0.9	3547	1	V21771	Expression vector
C 866	14	0.9	3553	1	O74208	VIR vaccination ve
C 867	14	0.9	3561	1	O85414	Human TYR2 kinase
C 868	14	0.9	3581	1	O64391	Sequence of murine
C 869	14	0.9	3582	1	V62430	Prostate cancer an
C 870	14	0.9	3597	1	V81414	Human prenylcysteil
C 871	14	0.9	3610	1	V00669	Synthetic hepariti
C 872	14	0.9	3631	1	O70733	TATa-binding prote
C 873	14	0.9	3651	1	T15906	HFRS virus strain
C 874	14	0.9	3656	1	T36751	VDL receptor gene
C 875	14	0.9	3660	1	T95249	M. catarrhalis Q8
C 876	14	0.9	3744	1	V23296	Streptococcus pneu
C 877	14	0.9	3780	1	T61864	c-abl gene. Antise
C 878	14	0.9	3786	1	T89300	Human ST receptor
C 879	14	0.9	3786	1	T97229	Human ST receptor
C 880	14	0.9	3812	1	T42219	Drosophila TATa-bi
C 881	14	0.9	3812	1	T79605	TATa-binding prote
C 882	14	0.9	3835	1	X32295	M. grisea PTH11 ge
C 883	14	0.9	3840	1	N92365	Sequence of maize
C 884	14	0.9	3853	1	O04776	Cloned gene encodi
C 885	14	0.9	3853	1	O15324	Beta-amylase pmm1
C 886	14	0.9	3854	1	O04910	Cloned sequence en
C 887	14	0.9	3854	1	T70493	cDNA encoding E-se
C 888	14	0.9	3854	1	T97610	Human E2A-binding
C 889	14	0.9	3861	1	O05979	Recombinant ACNPV-
C 890	14	0.9	3863	1	O06686	Endothelial cell-1
C 891	14	0.9	3888	1	O04549	Type I interferon
C 892	14	0.9	3901	1	O71190	PLVY genomic fragm
C 893	14	0.9	3909	1	X02981	Human IL-1ra BAC c
C 894	14	0.9	3958	1	O29230	The abf-A gene of
C 895	14	0.9	3993	1	O15153	Swine herpes virus
C 896	14	0.9	4001	1	X29857	Thermostable neapoi
C 897	14	0.9	4034	1	V31354	Trichoderma reesei
C 898	14	0.9	4044	1	O15154	Plasmid pHAS2-WCP
C 899	14	0.9	4080	1	T30652	P. gingivalis haem
C 900	14	0.9	4080	1	V58873	Haemagglutinin pro
C 901	14	0.9	4113	1	O06644	Recombinant ACNPV-
C 902	14	0.9	4134	1	T58545	Human phosphatidy
C 903	14	0.9	4136	1	T95368	Human phosphatidy
C 904	14	0.9	4137	1	T58546	Human phosphatidy
C 905	14	0.9	4140	1	T59799	Aureobasidin resis
C 906	14	0.9	4154	1	O79037	Human SREBP-1a cod
C 907	14	0.9	4219	1	O04786	Sequence encoding
C 908	14	0.9	4278	1	V52286	Streptococcus pneu
C 909	14	0.9	4283	1	T86449	DNA encoding hGH,
C 910	14	0.9	4283	1	T86610	Epidermal growth f
C 911	14	0.9	4300	1	O50468	Glycosyltransferas
C 912	14	0.9	4300	1	O50982	Aspergillus niger
C 913	14	0.9	4307	1	O10988	Lambda clone 101 c
C 914	14	0.9	4307	1	O26078	Lambda clone 101.
C 915	14	0.9	4313	1	V84546	Human secreted pro
C 916	14	0.9	4320	1	N90349	Transcription sequ
C 917	14	0.9	4328	1	V21684	Vector plasmid CMV
C 918	14	0.9	4328	1	X24808	Vector CMVkm2. Lip
C 919	14	0.9	4348	1	V29008	Parathyroid hormon
C 920	14	0.9	4353	1	O04787	Sequence encoding
C 921	14	0.9	4377	1	V23493	Pseudomonas lipO,
C 922	14	0.9	4393	1	O81456	Orotidine-5'-monop
C 923	14	0.9	4395	1	V65398	DNA encoding an an
C 924	14	0.9	4432	1	O74206	Expression vector
C 925	14	0.9	4432	1	T01113	Expression vector
C 926	14	0.9	4433	1	T95023	Arabidopsis thalia
C 927	14	0.9	4456	1	V81748	Human Alp encoding
C 928	14	0.9	4465	1	T42135	IC2-2 gene encodin
C 929	14	0.9	4483	1	V52496	Streptococcus pneu
C 930	14	0.9	4503	1	T64783	Human oxygen regul
C 931	14	0.9	4503	1	X13086	Enterococcus faeca
C 932	14	0.9	4506	1	O35899	Beta-ketothiolase
C 933	14	0.9	4530	1	T01585	Her-2/neu (ERBB2/c
C 934	14	0.9	4530	1	T71253	Human HER2 gene. P
C 935	14	0.9	4534	1	X24069	Human 53BP2 DNA. C
C 936	14	0.9	4550	1	O14936	Abelson Related Ge
C 937	14	0.9	4566	1	V41550	Nucleotide sequenc
C 938	14	0.9	4566	1	X01089	Human G-protein co
C 939	14	0.9	4568	1	T44039	Human G-protein re
C 940	14	0.9	4568	1	X01094	Human G-protein co
C 941	14	0.9	4696	1	X03041	Human IL-1ra BAC c
C 942	14	0.9	4729	1	O04043	Recombinant DNA fo
C 943	14	0.9	4732	1	O51730	Plasmid pRK5.1 for
C 944	14	0.9	4732	1	X15636	Nucleotide sequenc
C 945	14	0.9	4736	1	X23315	Mouse 1-alpha-Ohas
C 946	14	0.9	4746	1	T47874	K. pneumoniae diol
C 947	14	0.9	4762	1	T94052	Human multidrug re
C 948	14	0.9	4779	1	T58751	Nucleolar/endosoma
C 949	14	0.9	4788	1	O85367	MP-1-alpha. Scree
C 950	14	0.9	4818	1	V21686	Vector plasmid CMV
C 951	14	0.9	4864	1	O74207	Expression vector
C 952	14	0.9	4864	1	T01114	Expression vector
C 953	14	0.9	4864	1	V21737	Vibrio sequence us
C 954	14	0.9	4887	1	O41290	Ubiquitin-specific
C 955	14	0.9	4912	1	X08423	Recombinant vector
C 956	14	0.9	4928	1	T32656	Plasmid encoding h
C 957	14	0.9	4928	1	V70138	Pyruvate formate l
C 958	14	0.9	4937	1	O49507	Sequence encoding
C 959	14	0.9	4937	1	O72294	Haemophilus high m
C 960	14	0.9	4937	1	T90995	Non-typeable Haemo
C 961	14	0.9	4948	1	T42134	IC2-1 gene encodin
C 962	14	0.9	4963	1	O76270	PhCMV*-1-controlle
C 963	14	0.9	4963	1	T06873	PhCMV*-1 promoter
C 964	14	0.9	4963	1	X01370	PhCMV*-1 rabbit pr
C 965	14	0.9	4963	1	X27906	Rabbit progesteron
C 966	14	0.9	4992	1	N80297	Partial nucleic ac
C 967	14	0.9	5030	1	T75702	pgREENLANTERN-1 ve
C 968	14	0.9	5030	1	X08453	Plasmid pKlen-TRD1
C 969	14	0.9	5060	1	T78171	Plasmid pSE-2 enco
C 970	14	0.9	5062	1	T78170	Nucleotide sequenc
C 971	14	0.9	5090	1	V35375	Enterococcus faeca
C 972	14	0.9	5101	1	X13453	Vector plasmid CMV
C 973	14	0.9	5107	1	V21685	DNA encoding solub
C 974	14	0.9	5140	1	O24977	Plasmid pRK5c. Pro
C 975	14	0.9	5158	1	X28081	prct-splice. Nucle
C 976	14	0.9	5178	1	T49876	Mouse patched gene
C 977	14	0.9	5187	1	T14218	Mouse patched (ptc
C 978	14	0.9	5187	1	V21589	Precis coenia (but
C 979	14	0.9	5187	1	V21587	Mouse patched gene
C 980	14	0.9	5187	1	V64092	Precis coenia patc
C 981	14	0.9	5187	1	V64099	Plasmid nKCMVintBL
C 982	14	0.9	5218	1	V00682	

C 983	14	0.9	5288	1	T14220	Human patched gene
C 984	14	0.9	5288	1	V21590	Human patched (pic
C 985	14	0.9	5288	1	V64093	Human patched gene
C 986	14	0.9	5321	1	V72945	Rat Munc13-1 encod
C 987	14	0.9	5340	1	N71017	Sequence of LAV/HT
C 988	14	0.9	5471	1	T35874	Herpesvirus of tur
C 989	14	0.9	5496	1	Q73395	CV1J1 coding seque
C 990	14	0.9	5504	1	X20559	Polynucleotide seq
C 991	14	0.9	5558	1	V52236	Streptococcus pneu
C 992	14	0.9	5585	1	087418	Plasmd pmpc-112.
C 993	14	0.9	5585	1	T61430	Interleukin-2 expr
C 994	14	0.9	5620	1	T48630	Ad5-TFR, CMV-Lucif
C 995	14	0.9	5695	1	X20636	Polynucleotide seq
C 996	14	0.9	5706	1	Q13572	E75A transcription
C 997	14	0.9	5778	1	N40033	Sequence of portio
C 998	14	0.9	5820	1	T71205	hSNV polyhedrin
C 999	14	0.9	5821	1	T58686	DNA encoding S. fr
C1000	14	0.9	5843	1	V00676	DNA plasmd VR2210

ALIGNMENTS

```
RESULT 1
X00070 standard: DNA; 1491 BP.
AC 17-MAR-1999 (first entry)
DE Aspergillus oryzae dipeptidyl aminopeptidase encoding DNA #2.
KW Prolyl dipeptidyl aminopeptidase; protein hydrolyase; dough;
KM flavour enhancer; palatability; mouthfeel; aroma; crust colour;
KW baking; animal feed additive; hydrolysis; ss.
OS Aspergillus oryzae.
FH key Location/Qualifiers
FT CDS 1..1491
FT     /*tag= a
PN NC09851803-A1.
PF 19-NOV-1998.
PR 12-MAY-1998: U09629.
PR 20-OCT-1997: US-062892.
PR 16-MAY-1997: US-857884.
PA (NOVO ) NOVO NORDISK BIOTECH INC.
PI Blinkovsky A, Brown K, Byun T, Klotz A, Rey MW;
PI WPI: 99-045322/04.
DR P-SDB: M89614.
PT New dipeptidyl aminopeptidase from Aspergillus oryzae - used to
PT produce protein hydrolysates enriched in particular amino acids,
PT useful as flavour enhancers, e.g. in doughs
PS Claim 1: Fig 1: 77pp: English.
CC The present sequence encodes dipeptidyl aminopeptidase (DPAP) from
CC Aspergillus oryzae. DPAP acts synergistically with an aminopeptidase
CC (AP) to hydrolyse polypeptides, producing protein hydrolysate (PH),
CC useful in foods as flavour enhancer, e.g. in baked goods, enriched in:
CC (a) Ala, Arg, Asp, Gly and/or Val, or (b), if the substrate has been
CC deamidated, in Glu (free and/or peptide bound), in which case products
CC are useful as animal feed additives. DPAP can also be used in flavour-
CC improving compositions (optionally containing AP) and in dough pre-mixes,
CC also for desactivating enzymes and for converting precursors to mature
CC proteins. DPAP increases the level of hydrolysis of proteins and thus
CC of flavour development, and a mixture with AP may hydrolyse tripeptides
CC that are resistant to either enzyme used alone. PH have improved
CC solubility, emulsifying and foaming properties, and products containing
CC them have better flavour, palatability, mouthfeel, aroma and crust
CC colour.
SQ Sequence 1491 BP; 335 A; 438 C; 426 G; 292 T;
```

```
Query Match 100.0%: Score 1491; DB 1: Length 1491;
Best Local Similarity 100.0%: Pred. No. 0;
Matches 1491; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGAGTGCCTTTGTGGGCTTCGTTGCGGCGTGTGGCTGGAGGCGCTTGT 60
|||||
```

```
Db 1 ATGAGTGCCTTTGTGGGCTTCGTTGCGGCGTGTGGCTGGAGGCGCTTGT 60
QY 1 TCGCGGATGAGTCCCGGAGATATTCAGTTGGAAATCTGCTGGAAAGATCCCAAG 120
61 |||||||
Db 1 TCGCGGATGAGTCCCGGAGATATTCAGTTGGAAATCTGCTGGAAAGATCCCAAG 120
QY 121 CTTGAGACTTCGCGCTTATGCTTACCCGAGCGCAATGCGCTCTTGGTGAAGCCAC 180
121 |||||||
Db 121 CTTGAGACTTCGCGCTTATGCTTACCCGAGCGCAATGCGCTCTTGGTGAAGCCAC 180
QY 121 CTTGAGACTTCGCGCTTATGCTTACCCGAGCGCAATGCGCTCTTGGTGAAGCCAC 180
121 CTTGAGACTTCGCGCTTATGCTTACCCGAGCGCAATGCGCTCTTGGTGAAGCCAC 180
QY 181 GAGCAGACGGTTAACTATCTCTACGAGAGCTGAAGAAGCTGGCTATATGATGTAC 240
181 |||||||
Db 181 GAGCAGACGGTTAACTATCTCTACGAGAGCTGAAGAAGCTGGCTATATGATGTAC 240
QY 181 GAGCAGACGGTTAACTATCTCTACGAGAGCTGAAGAAGCTGGCTATATGATGTAC 240
181 GAGCAGACGGTTAACTATCTCTACGAGAGCTGAAGAAGCTGGCTATATGATGTAC 240
QY 241 AAGAGCCTCAGTGCACCTGTGAGCAATGCCAGACCGTCAAGGTCGAGAGAG 300
241 |||||||
Db 241 AAGAGCCTCAGTGCACCTGTGAGCAATGCCAGACCGTCAAGGTCGAGAGAG 300
QY 241 AAGAGCCTCAGTGCACCTGTGAGCAATGCCAGACCGTCAAGGTCGAGAGAG 300
241 AAGAGCCTCAGTGCACCTGTGAGCAATGCCAGACCGTCAAGGTCGAGAGAG 300
QY 301 GAAATCGAGGCGAAGACCATGACCTACAGTCCAGCGTGCAGGCTGAGTATGCC 360
301 |||||||
Db 301 GAAATCGAGGCGAAGACCATGACCTACAGTCCAGCGTGCAGGCTGAGTATGCC 360
QY 361 GTGCTCAAGAACTGGATGCGAGCGAGCGGATTAACCATCCGATGTGAGGCAAGTGC 420
361 |||||||
Db 361 GTGCTCAAGAACTGGATGCGAGCGAGCGGATTAACCATCCGATGTGAGGCAAGTGC 420
QY 421 GCCCTGATCAAGGCTGGAGAAATGCCCTGGCGGACAAAGTGGTTCGCTGCAAAAGCC 480
421 |||||||
Db 421 GCCCTGATCAAGGCTGGAGAAATGCCCTGGCGGACAAAGTGGTTCGCTGCAAAAGCC 480
QY 481 AAGGCGCGGCTTCGATGCTATTAACAATGTGGCGCGATGCATGCGGCGACCTTGGC 540
481 |||||||
Db 481 AAGGCGCGGCTTCGATGCTATTAACAATGTGGCGCGATGCATGCGGCGACCTTGGC 540
QY 481 AAGGCGCGGCTTCGATGCTATTAACAATGTGGCGCGATGCATGCGGCGACCTTGGC 540
481 AAGGCGCGGCTTCGATGCTATTAACAATGTGGCGCGATGCATGCGGCGACCTTGGC 540
QY 541 GCGGCGCAGAGTAAAGGACCGTATTCGCGCATTTGCGTATCAGTTGAGAGATGCC 600
541 |||||||
Db 541 GCGGCGCAGAGTAAAGGACCGTATTCGCGCATTTGCGTATCAGTTGAGAGATGCC 600
QY 541 GCGGCGCAGAGTAAAGGACCGTATTCGCGCATTTGCGTATCAGTTGAGAGATGCC 600
541 GCGGCGCAGAGTAAAGGACCGTATTCGCGCATTTGCGTATCAGTTGAGAGATGCC 600
QY 601 CAGAGCTGATCAAGCTTGTGCTGAGAGCTGAGTATCTGTGATCTGTGGTGGATAGT 660
601 |||||||
Db 601 CAGAGCTGATCAAGCTTGTGCTGAGAGCTGAGTATCTGTGATCTGTGGTGGATAGT 660
QY 601 CAGAGCTGATCAAGCTTGTGCTGAGAGCTGAGTATCTGTGATCTGTGGTGGATAGT 660
601 CAGAGCTGATCAAGCTTGTGCTGAGAGCTGAGTATCTGTGATCTGTGGTGGATAGT 660
QY 661 AAGCAGAGAAACGCTACGAGATATACCTTGTGCGGACGAGAGAGGCGCGATCCGAAC 720
661 |||||||
Db 661 AAGCAGAGAAACGCTACGAGATATACCTTGTGCGGACGAGAGAGGCGCGATCCGAAC 720
QY 721 AACGCTGCGCGTGGGTGACACGAGCTCAAGTCAAGGCGGCGCTGTATCAAGAC 780
721 |||||||
Db 721 AACGCTGCGCGTGGGTGACACGAGCTCAAGTCAAGGCGGCGCTGTATCAAGAC 780
QY 781 GATGCTCGGGGCAATATTAACCACTTGTGATTCGCAAAAGGCTCAGCGTCCGCTC 840
781 |||||||
Db 781 GATGCTCGGGGCAATATTAACCACTTGTGATTCGCAAAAGGCTCAGCGTCCGCTC 840
QY 841 AAGATATCCCTGCGCTTCCTTCTTGACAGCAGAGAGAGTTCGCTGCGGACAGAAC 900
841 |||||||
Db 841 AAGATATCCCTGCGCTTCCTTCTTGACAGCAGAGAGAGTTCGCTGCGGACAGAAC 900
QY 841 AAGATATCCCTGCGCTTCCTTCTTGACAGCAGAGAGAGTTCGCTGCGGACAGAAC 900
841 AAGATATCCCTGCGCTTCCTTCTTGACAGCAGAGAGAGTTCGCTGCGGACAGAAC 900
QY 901 TACTACTCTCCATCTGAATGCGACCGAGCTGAACAAGTCCGACTGTACGAACTTC 960
901 |||||||
Db 901 TACTACTCTCCATCTGAATGCGACCGAGCTGAACAAGTCCGACTGTACGAACTTC 960
QY 901 TACTACTCTCCATCTGAATGCGACCGAGCTGAACAAGTCCGACTGTACGAACTTC 960
901 TACTACTCTCCATCTGAATGCGACCGAGCTGAACAAGTCCGACTGTACGAACTTC 960
QY 961 GACATGATGCGCTCAGTACTAGCGCCCTATGATATATAGCGTATGATGCGGCTTC 1020
961 |||||||
Db 961 GACATGATGCGCTCAGTACTAGCGCCCTATGATATATAGCGTATGATGCGGCTTC 1020
QY 961 GACATGATGCGCTCAGTACTAGCGCCCTATGATATATAGCGTATGATGCGGCTTC 1020
961 GACATGATGCGCTCAGTACTAGCGCCCTATGATATATAGCGTATGATGCGGCTTC 1020
QY 1021 AACGAGAGGAGGCGGCGGCTTCGCGCAGATGAGAACTGTTGAGAGCTACTAGAC 1080
1021 |||||||
Db 1021 AACGAGAGGAGGCGGCGGCTTCGCGCAGATGAGAACTGTTGAGAGCTACTAGAC 1080
QY 1021 AACGAGAGGAGGCGGCGGCTTCGCGCAGATGAGAACTGTTGAGAGCTACTAGAC 1080
1021 AACGAGAGGAGGCGGCGGCTTCGCGCAGATGAGAACTGTTGAGAGCTACTAGAC 1080
QY 1081 TCCATGACCTGCGCTATATCCCAACCACTTTGAGGAGAGTTCGCGACTGAGGCTTT 1140
1081 |||||||
Db 1081 TCCATGACCTGCGCTATATCCCAACCACTTTGAGGAGAGTTCGCGACTGAGGCTTT 1140
```

QY 1141 ATCTGAACGGATCTCCCTGCGTGTGACTTTCAAGGGCGCGAGGATCATGTCGGA 1200
 |||||
 Db 1141 ATCTGAACGGATCTCCCTGCGTGTGACTTTCAAGGGCGCGAGGATCATGTCGGA 1200
 QY 1201 GAGAACGCAAGCCGCTGGGAGGAGTCAAGCCGCGGTGACCTAGAGCCCACTACAGCC 1260
 |||||
 Db 1201 GAGAACGCAAGCCGCTGGGAGGAGTCAAGCCGCGGTGACCTAGAGCCCACTACAGCC 1260
 QY 1261 GCGGAGACAACTATGACCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1320
 |||||
 Db 1261 GCGGAGACAACTATGACCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1320
 QY 1321 GCGTTCGCGTGGCCACTTACGCGCAAGACCTTCCTGATCCCAAGGATACGACA 1380
 |||||
 Db 1321 GCGTTCGCGTGGCCACTTACGCGCAAGACCTTCCTGATCCCAAGGATACGACA 1380
 QY 1381 TCTCTCTTGCACCGACGAGCCCGCACATGCGACCATTCGCGAAGAGAGCTCCGAAGACA 1440
 |||||
 Db 1381 TCTCTCTTGCACCGACGAGCCCGCACATGCGACCATTCGCGAAGAGAGCTCCGAAGACA 1440
 QY 1441 CACGCTACGATACGATCGGATGCTGCAATTCCTCAAGTCGAGGATAG 1491
 |||||
 Db 1441 CACGCTACGATACGATCGGATGCTGCAATTCCTCAAGTCGAGGATAG 1491

RESULT 2
 V82521
 ID V82521 standard; DNA: 1491 BP.
 AC V82521;
 DT 17-MAR-1999 (first entry)
 DE Aspergillus oryzae aminopeptidase II encoding DNA.
 KW Aminopeptidase; protein hydrolyase; glycine releasing; protease;
 KW proteolaceous material; flavour; food; baking; animal feed additive;
 KW palatable; hydrolysis; solubility; emulsifying; foaming; aroma;
 KW mouthfeel; crust colour; ss.
 OS Aspergillus oryzae.
 FH Key Location/Qualifiers
 FT CDS 1..1491
 FT 1.1491
 FT /*tag= a
 PN M09851163-A2.
 PD 19-NOV-1998.
 PE 15-MAY-1998; U09998.
 PR 16-DEC-1997; US-069719.
 PR 16-MAY-1997; US-857886.
 PR 20-OCT-1997; US-062893.
 PR 16-DEC-1997; DK-001465.
 PA (ASAH) ASAH CHEM IND CO LTD.
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 PA (NOVO) NOVO NORDISK AS.
 PI Blinovsky A, Brown K, Byun T, Fujii M, Golightly E,
 PI Kotod LV, Marumotac, Mathiasen TE;
 DR P-PSDB; W89586.
 DR WPI; 99-045177/04.
 PT Production of protein hydrolysate - using protease and enzyme that
 PT releases glycine, useful as flavour improvers in foods and animal
 PT feed additives
 PS Claim 4; Fig 1; 84pp. English.
 CC A method has been developed for the production of protein hydrolysates
 CC (PH) comprising reacting a protein with: (i) at least one polypeptide
 CC with Gly-releasing activity; and (ii) at least one other protease so
 CC that the amount of Gly produced is greater than when the protease is
 CC used alone. PH are used to improve flavour of foods (e.g. baked goods)
 CC and as animal feed additives. PH are preferably also enriched in Glu
 CC (free and/or peptide bound), so have improved flavour and palatability.
 CC Addition of a polypeptide with Gly-releasing activity increases the
 CC degree of hydrolysis (or reduces the amount of enzyme needed) and
 CC hydrolysates have better solubility, and emulsifying and foaming
 CC properties. Baked goods containing them have improved aroma, mouthfeel
 CC and crust colour. The present sequence encodes Aspergillus oryzae
 CC aminopeptidase II, which is used in the method of the invention.
 SQ Sequence 1491 BP; 335 A; 438 C; 426 G; 292 T;

Query Match 100.0%; Score 1491; DB 1; Length 1491;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1491; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGAGCTGCTTTTGTGGCTTCTGTTGCTTGGGCGTGTGGGAGGAGCGCTTGT 60
 |||||
 Db 1 ATGAGCTGCTTTTGTGGCTTCTGTTGCTTGGGCGTGTGGGAGGAGCGCTTGT 60
 QY 61 TCGCCGATGAGTTTCCCGAGAGATTTGATGGAGATCTGCTGGAAGATCCCAAC 120
 |||||
 Db 61 TCGCCGATGAGTTTCCCGAGAGATTTGATGGAGATCTGCTGGAAGATCCCAAC 120
 QY 121 CTGAGGACTTCCCTATGCTTACCCCGAGGCAATCCGCTTGTGGTAAAGCCAC 180
 |||||
 Db 121 CTGAGGACTTCCCTATGCTTACCCCGAGGCAATCCGCTTGTGGTAAAGCCAC 180
 QY 181 GACGACAGGTTTAACTATCTACGAGAGGCTGAAGAAGACTGCTACTATGATGTAC 240
 |||||
 Db 181 GACGACAGGTTTAACTATCTACGAGAGGCTGAAGAAGACTGCTACTATGATGTAC 240
 QY 241 AAGCAGCTCAGGTGACCTGTGAGCAATGCCGACAGAGCTTCMAAGTGGCGATAG 300
 |||||
 Db 241 AAGCAGCTCAGGTGACCTGTGAGCAATGCCGACAGAGCTTCMAAGTGGCGATAG 300
 QY 301 GAAATCGAGCGCAAGACCATGACCTACAGTCCCAAGCTTCAGAGTACCGCGATGACC 360
 |||||
 Db 301 GAAATCGAGCGCAAGACCATGACCTACAGTCCCAAGCTTCAGAGTACCGCGATGACC 360
 QY 361 GTCTCAAGAACTTGGATGACGAGCGAGGATTTACCATCGATGTGAGAGGCAAGTTC 420
 |||||
 Db 361 GTCTCAAGAACTTGGATGACGAGCGAGGATTTACCATCGATGTGAGAGGCAAGTTC 420
 QY 421 GCGCTGATCAAGCTGGAAGATGCCGTTGCGGCAAGATCGGTTCTGCTGCCAAAGCC 480
 |||||
 Db 421 GCGCTGATCAAGCTGGAAGATGCCGTTGCGGCAAGATCGGTTCTGCTGCCAAAGCC 480
 QY 481 AAGCGCGGCGTTCGATGTTCTATTAACAATGCGCGGATCCATGCGCGGACACCTTGGC 540
 |||||
 Db 481 AAGCGCGGCGTTCGATGTTCTATTAACAATGCGCGGATCCATGCGCGGACACCTTGGC 540
 QY 541 GCGGCGCAGAGTAAAGGACCGGATTCGCGCATTTGCGGATCAGTGTGAGAGATGCG 600
 |||||
 Db 541 GCGGCGCAGAGTAAAGGACCGGATTCGCGCATTTGCGGATCAGTGTGAGAGATGCG 600
 QY 601 CAGAACTGATCAAGCTTCTGAGGCTGGATCGATCTGTGATCTGTGGGTGATAGT 660
 |||||
 Db 601 CAGAACTGATCAAGCTTCTGAGGCTGGATCGATCTGTGATCTGTGGGTGATAGT 660
 QY 661 AAGCAGGAGAACCTGACAGCATATTAAGCTTGTCCGCGCAGAGAGGCGCGATCCGAC 720
 |||||
 Db 661 AAGCAGGAGAACCTGACAGCATATTAAGCTTGTCCGCGCAGAGAGGCGCGATCCGAC 720
 QY 721 AAGCTGCTGCGGCTGGGAGGCGCAGAGCTACGTCGAGGCGGCGCTTGATCAAGAC 780
 |||||
 Db 721 AAGCTGCTGCGGCTGGGAGGCGCAGAGCTACGTCGAGGCGGCGCTTGATCAAGAC 780
 QY 781 GATGCTCGGCGCATTAATAGCAATTTGGTCAATTTGCCAAGGCGCTCAGCAGATCCGTC 840
 |||||
 Db 781 GATGCTCGGCGCATTAATAGCAATTTGGTCAATTTGCCAAGGCGCTCAGCAGATCCGTC 840
 QY 841 AAGATGCGGTGCGTCTCTCTTCTGACAGCAGAGAGGTTGCTGCTGGGACGAC 900
 |||||
 Db 841 AAGATGCGGTGCGTCTCTCTTCTGACAGCAGAGAGGTTGCTGCTGGGACGAC 900
 QY 901 TACTAGCTCTCCATCTGAATGCCAGGAGTGAACAAGATCCGACTTACTGAACTTC 960
 |||||
 Db 901 TACTAGCTCTCCATCTGAATGCCAGGAGTGAACAAGATCCGACTTACTGAACTTC 960
 QY 961 GACATGATGCGCTCAGTAACTAGCGCTCATGATCTATGACGATGATGCGGCTTC 1020
 |||||
 Db 961 GACATGATGCGCTCAGTAACTAGCGCTCATGATCTATGACGATGATGCGGCTTC 1020

QY 1021 AACGAGCGGACCGCGCGTTCGCCGAGATCGAAGAACTGTCGAGAGCTACTACGAC 1080
|||||
Db 1021 AACGAGCGGACCGCGCGTTCGCCGAGATCGAAGAACTGTCGAGAGCTACTACGAC 1080
QY 1081 TCCATCGACCTGCGCTCATATCCCAAGTTTGACGAGAGCTTCGAGTACGAGCCCTTT 1140
|||||
Db 1081 TCCATCGACCTGCGCTCATATCCCAAGTTTGACGAGAGCTTCGAGTACGAGCCCTTT 1140
QY 1141 ATCTGAGAGCGGATTCCTGCGGTGAGACTTTCACGCGCGCGAGAGGATCATGTCGAA 1200
|||||
Db 1141 ATCTGAGAGCGGATTCCTGCGGTGAGACTTTCACGCGCGCGAGAGGATCATGTCGAA 1200
QY 1201 GAGAACGCAACCCGCTGGGAGGAGTCAACCGCGCGTGGCTCAGAGCCCACTACACGCC 1260
|||||
Db 1201 GAGAACGCAACCCGCTGGGAGGAGTCAACCGCGCGTGGCTCAGAGCCCACTACACGCC 1260
QY 1261 GCGGGAGACAACTGACCAACCTCAGACCTTCCTGATCACTCCAAAGCCAC 1320
|||||
Db 1261 GCGGGAGACAACTGACCAACCTCAGACCTTCCTGATCACTCCAAAGCCAC 1320
QY 1321 GCGTTGCGCGCTGCGCACTACGCGCAAGACCTTCCTGATCCCAAGGAAATACAC 1380
|||||
Db 1321 GCGTTGCGCGCTGCGCACTACGCGCAAGACCTTCCTGATCCCAAGGAAATACAC 1380
QY 1381 TCCCTGCTGACGACGACGACCCGACCATGCGACCATTCGCGCAAGAGAGCTCCGAAAGCA 1440
|||||
Db 1381 TCCCTGCTGACGACGACGACCCGACCATGCGACCATTCGCGCAAGAGAGCTCCGAAAGCA 1440
QY 1441 CACGCTCAGCTATCAGATCCGGATGCTGCGATTCCTCAAGTCGAGGCTAG 1491
|||||
Db 1441 CACGCTCAGCTATCAGATCCGGATGCTGCGATTCCTCAAGTCGAGGCTAG 1491
RESULT 3
X00018
ID X00018 standard; DNA; 1491 BP.
AC X00018;
DT 22-MAR-1999 (first entry)
DE Aspergillus oryzae aminopeptidase II encoding DNA.
KW aminopeptidase; protein hydrolyase; baking; dough; flavour enhancer;
KM mouthfeel; palatability; aroma; hydrolysis; animal feed additive; ss.
OS Aspergillus oryzae.
FH Key Location/Qualifiers
FT CDS 1..1491
FT /note- "a
FT /note- "at position 277 to 279 the DNA encodes Glu
FT (position 93 in the protein) which corresponds
FT to the protein given in the sequence listing,
FT but the protein given in the figure has Asp at
FT this position"
PN MO9851804-A1.
PI 19-NOV-1998
PD 15-MAY-1998; U09940.
PR 20-OCT-1997; US-062893.
PR 16-MAY-1997; US-857886.
PA (NOVO) NOVO NORDISK BIOTECH INC.
PI Blinkovsky A, Brown K, Byun T, Gollightly E, Kofod LV;
DR WPI: 99-045233/04.
DR P-PSDB: W89597.
PT New aminopeptidase from Aspergillus oryzae - used to produce protein
PT enhancers, e.g. in doughs
PT Claim 1: Fig 1: 92pp: English.
PS The present sequence encodes aminopeptidase (AP) II from Aspergillus
CC oryzae. AP is used in combination with an endopeptidase (EP) to
CC hydrolyse polypeptides, producing protein hydrolysate (PH), useful in
CC foods as flavour enhancer, e.g. in baked goods, enriched in: (a) Leu,
CC Gly (especially), Glu, Ser, Asp, Asn, Pro, Cys, Ala and/or Glu, or (b),
CC if the substrate has been deamidated, in Glu (free and/or peptide bound),
CC in which case products are used as animal feed additives. AP can be
CC used in flavour-improving compositions (optionally containing EP) and in
CC dough pre-mixes. Also AP can be used for deactivating enzymes and for
CC converting precursors to mature proteins. AP increases the level of

CC hydrolysis of proteins and thus of flavour development. PH have improved
CC solubility, emulsifying and foaming properties, and products containing
CC them have better flavour, palatability and aroma.
S0 Sequence 1491 BP; 335 A; 438 C; 426 G; 292 T;
Query Match 100.0%; Score 1491; DB 1; Length 1491;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1491; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGTTCCTTTTGTGGGCTTCGTTGCGGCGCTGTGGCTGGAGGCGCTTGT 60
Db 1 ATGAGTTCCTTTTGTGGGCTTCGTTGCGGCGCTGTGGCTGGAGGCGCTTGT 60
QY 61 TCGCGGAGTGAATCCCGAGAGATTCAGTTGGAAGATCTGCTGGAAGATCCCAAC 120
Db 61 TCGCGGAGTGAATCCCGAGAGATTCAGTTGGAAGATCTGCTGGAAGATCCCAAC 120
QY 121 CTTGAGAGCTTCGCTTATGCTTACCCGAGCGCAATCGCTTGGTGAAGCCAC 180
Db 121 CTTGAGAGCTTCGCTTATGCTTACCCGAGCGCAATCGCTTGGTGAAGCCAC 180
QY 181 GACGACACGTTAACTATCTCTACGAGAGCTGAAGAAGACTGCTACTATGATGTCTAC 240
Db 181 GACGACACGTTAACTATCTCTACGAGAGCTGAAGAAGACTGCTACTATGATGTCTAC 240
QY 241 AAGCAGCTCAGTGTGACCTGTGAGAGCAATCCGACAGAGCTCAAGGTGGCGATGAG 300
Db 241 AAGCAGCTCAGTGTGACCTGTGAGAGCAATCCGACAGAGCTCAAGGTGGCGATGAG 300
QY 301 GAAATCGAGCGAAGACCATATGCTTACAGTCCAGCGTCCAGGCTCAGCGCATGTAGCC 360
Db 301 GAAATCGAGCGAAGACCATATGCTTACAGTCCAGCGTCCAGGCTCAGCGCATGTAGCC 360
QY 361 GTCTCAAGAACTTGGATGTCAGGAGCGGATTAACCATCCGATGTCGAGGCAAGTGC 420
Db 361 GTCTCAAGAACTTGGATGTCAGGAGCGGATTAACCATCCGATGTCGAGGCAAGTGC 420
QY 421 GCGCTGATCAAGCGGTGAGATGCGCGTGGCGAGCAAGTGGTTCGCTGCCAAAGCC 480
Db 421 GCGCTGATCAAGCGGTGAGATGCGCGTGGCGAGCAAGTGGTTCGCTGCCAAAGCC 480
QY 481 AAGCGCGGCGCTTGTATGATGATTAACAAATGTCGCGGATCCATGAGGCGACCTTGGC 540
Db 481 AAGCGCGGCGCTTGTATGATGATTAACAAATGTCGCGGATCCATGAGGCGACCTTGGC 540
QY 541 GCGCGCAGAGATGTAAGAGGACCGTATTCGCGCATTTGCGTATCAGCTTGAAGATGCG 600
Db 541 GCGCGCAGAGATGTAAGAGGACCGTATTCGCGCATTTGCGTATCAGCTTGAAGATGCG 600
QY 601 CAGAAAGTGAATCAAGCTTGTGAGAGCGTGAATCGGTATCTGTGGATCTGTGGTGGATAGT 660
Db 601 CAGAAAGTGAATCAAGCTTGTGAGAGCGTGAATCGGTATCTGTGGATCTGTGGTGGATAGT 660
QY 661 AAGCAGAGAAACCTTACGACGTATTAACGTTGTCGCGAGAGAAAGGCGCGCATCCGAC 720
Db 661 AAGCAGAGAAACCTTACGACGTATTAACGTTGTCGCGAGAGAAAGGCGCGCATCCGAC 720
QY 721 AAGCTGCTCGCGGTGGGTGGCCACAGCACTCACTGAGGCGGCGCTGTATCAACGAC 780
Db 721 AAGCTGCTCGCGGTGGGTGGCCACAGCACTCACTGAGGCGGCGCTGTATCAACGAC 780
QY 781 GATGGCTCGGCGATTAATAGCACTTGTGATTCGCAAGAGCGCTCAGCGATCTCCGTC 840
Db 781 GATGGCTCGGCGATTAATAGCACTTGTGATTCGCAAGAGCGCTCAGCGATCTCCGTC 840
QY 841 AAGAAATCCGTCCTCTCTTGTGACAGCAGAGAGAGTTCGTCGAGGAGCAAC 900
Db 841 AAGAAATCCGTCCTCTCTTGTGACAGCAGAGAGAGTTCGTCGAGGAGCAAC 900
QY 901 TACTAGCTTCCATCTGAATGCCACGAGCTGAACAAGATCCGACTGTACTGAATCTTC 960
Db 901 TACTAGCTTCCATCTGAATGCCACGAGCTGAACAAGATCCGACTGTACTGAATCTTC 960

```
OY 961 GACATGATGCGCTCAGCTACGAGCCCTCATGATATGAGGAGATGCGCGCTTC 1020
DB 961 GACATGATGCGCTCAGCTACGAGCCCTCATGATATGAGGAGATGCGCGCTTC 1020
OY 1021 AACCAAGAGCGAGCGCGCGGTTCGCCAGATCGAAGAACTGTCGAGAGTACAC 1080
DB 1021 AACCAAGAGCGAGCGCGCGGTTCGCCAGATCGAAGAACTGTCGAGAGTACAC 1080
OY 1081 TCCATGACGCTGCTCATATATCCCAACCCAGTTTGACGAGCTTCGAGTACGAGSCTTT 1140
DB 1081 TCCATGACGCTGCTCATATATCCCAACCCAGTTTGACGAGCTTCGAGTACGAGSCTTT 1140
OY 1141 ATCTGTAAGCGGCTATCCCTCGCGTGGAGACTCTTTCAGGCGCGGAGGAGTATGTCGCA 1200
DB 1141 ATCTGTAAGCGGCTATCCCTCGCGTGGAGACTCTTTCAGGCGCGGAGGAGTATGTCGCA 1200
OY 1201 GAGAAAGCAAGCGCGCTGGGAGGATCAAGCCGCGGCTGACGAGCCAGTACACGCGCC 1260
DB 1201 GAGAAAGCAAGCGCGCTGGGAGGATCAAGCCGCGGCTGACGAGCCAGTACACGCGCC 1260
OY 1261 GCGGAGACAAATGACACCACTCAACCTGAAGCCTTCCTGATCACTCCAAAGCCACCC 1320
DB 1261 GCGGAGACAAATGACACCACTCAACCTGAAGCCTTCCTGATCACTCCAAAGCCACCC 1320
OY 1321 GCGTTCGCGCGCTGCGCCAGTACGAGCCAGTCTCTCTGATGCCAAAGGAAATGACACA 1380
DB 1321 GCGTTCGCGCGCTGCGCCAGTACGAGCCAGTCTCTCTGATGCCAAAGGAAATGACACA 1380
OY 1381 TCGTCTTGGACCGAGCGAGCCGCGACATGAGACATTCGCGCAAGAGAGTCCGAAAGACA 1440
DB 1381 TCGTCTTGGACCGAGCGAGCCGCGACATGAGACATTCGCGCAAGAGAGTCCGAAAGACA 1440
OY 1441 CACGCTACGATACAGATCCGGATGCTGCGATTCCTCAAGTCGAGGACATAG 1491
DB 1441 CACGCTACGATACAGATCCGGATGCTGCGATTCCTCAAGTCGAGGACATAG 1491

RESULT 4
X00024/c
ID X00024 standard; DNA; 32 BP.
AC X00024;
DE 17-MAR-1999 (first entry)
DB Aspergillus oryzae aminopeptidase II PCR reverse primer #3.
KW Aminopeptidase; protein hydrolysate; baking; dough; flavour enhancer;
KW mouthfeel; palatability; aroma; hydrolysis; animal feed additive;
KW PCR primer; ss.
OS Synthetic.
DB Aspergillus oryzae.
PN W03851804-A1.
PD 15-NOV-1998.
PE 15-MAY-1998; U099940.
PR 20-OCT-1997; US-062893.
PR 16-MAY-1997; US-857886.
PA (NOVO ) NOVO NORDISK BIOTECH INC.
PI Blinkovsky A, Brown K, Bynun T, Golightly E, Kofod LV;
DR WPI; 99-045323/04.
PT New aminopeptidase from Aspergillus oryzae - used to produce protein
PT hydrolysates enriched in particular amino acids, used as flavour
PT enhancers, e.g. in doughs
PS Example 12; Page 55; 92pp; English.
CC The present sequence represents a PCR primer for aminopeptidase
CC (AP) II. AP is used in combination with an endopeptidase (EP) to
CC hydrolyse polypeptides, producing protein hydrolysate (PH), useful in
CC foods as flavour enhancer, e.g. in baked goods, enriched in: (a) Ieu,
CC Gly (especially), Glu, Ser, Asp, Asn, Pro, Cys, Ala and/or Gln, or (b),
CC if the substrate has been deamidated, in Glu (free and/or peptide bound),
CC in which case products are useful as animal feed additives. AP can be
CC used in flavour-improving compositions (optionally containing EP) and in
CC dough pre-mixes. Also AP can be used for deactivating enzymes and for
CC converting precursors to mature proteins. AP increases the level of
CC hydrolysis of proteins and thus of flavour development. PH have improved
CC solubility, emulsifying and foaming properties, and products containing
```

```
CC them have better flavour, palatability and aroma.
SQ Sequence 32 BP; 9 A; 5 C; 8 G; 10 T;

Query Match 1.4%; Score 21; DB 1; Length 32;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1471 CATTCCTCAAGTCGAGGACATAG 1491
DB 32 CATTCCTCAAGTCGAGGACATAG 12

RESULT 5
X00012/c
ID X00012 standard; DNA; 32 BP.
AC X00012;
DE 17-MAR-1999 (first entry)
DB Aspergillus oryzae aminopeptidase II PCR reverse primer #2.
KW Aminopeptidase; protein hydrolysate; glycine releasing; protease;
KW proteinaceous material; flavour; food; baking; animal feed additive;
KW palatability; hydrolysis; solubility; emulsifying; foaming; aroma;
KW mouthfeel; crust colour; PCR primer; ss.
OS Synthetic.
DB Aspergillus oryzae.
PN W0385163-A2.
PD 15-NOV-1998.
PE 15-MAY-1998; U09998.
PR 16-DEC-1997; US-069719.
PR 16-MAY-1997; US-857886.
PR 20-OCT-1997; US-062893.
PR 16-DEC-1997; DK-001465.
PA (ASAHI ) ASAHI CHEM IND CO LTD.
PA (NOVO ) NOVO NORDISK BIOTECH INC.
PI Blinkovsky A, Brown K, Bynun T, Fujii M, Golightly E,
PI Kofod LV, Marumotac, Mathiansen TE;
DR WPI; 99-045177/04.
PT Production of protein hydrolysate - using protease and enzyme that
PT releases glycine, useful as flavour improvers in foods and animal
PT feed additives
PS Example 12; Page 28; 84pp; English.
CC A method has been developed for the production of protein hydrolysates
CC (PH) comprising reacting a protein with: (i) at least one polypeptide
CC with gly-releasing activity; and (ii) at least one other protease so
CC that the amount of Gly produced is greater than when the protease is
CC used alone. PH are used to improve flavour of foods (e.g. baked goods)
CC and as animal feed additives. PH are preferably also enriched in Glu
CC (free and/or peptide bound), so have improved flavour and palatability.
CC Addition of a polypeptide with gly-releasing activity increases the
CC degree of hydrolysis (or reduces the amount of enzyme needed) and
CC hydrolysates have better solubility, and emulsifying and foaming
CC properties. Baked goods containing them have improved aroma, mouthfeel
CC and crust colour. The present sequence represents a PCR primer for
CC Aspergillus oryzae aminopeptidase II, which is used in the method of
CC the invention.
SQ Sequence 32 BP; 9 A; 5 C; 8 G; 10 T;

Query Match 1.4%; Score 21; DB 1; Length 32;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1471 CATTCCTCAAGTCGAGGACATAG 1491
DB 32 CATTCCTCAAGTCGAGGACATAG 12

RESULT 6
X00021
ID X00021 standard; DNA; 23 BP.
AC X00021;
DE 17-MAR-1999 (first entry)
```

DE Aspergillus oryzae aminopeptidase II PCR forward primer #1.
 KM Aminopeptidase; protein hydrolysate; baking; dough; flavour enhancer;
 KM mouthfeel; palatability; aroma; hydrolysis; animal feed additive;
 KM PCR primer; ss.
 OS Synthetic.
 PN Aspergillus oryzae.
 PN WO9851804-A1.
 PN 19-NOV-1998.
 PF 15-MAY-1998: U09940.
 PR 20-OCT-1997: US-062893.
 PR 16-MAY-1997: US-857886.
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 PI Blinovsky A, Brown K, Byun T, Golightly E, Kofod LV;
 DR WPI: 99-045233/04.
 PT New aminopeptidase from Aspergillus oryzae - used to produce protein
 PT hydrolysates enriched in particular amino acids, used as flavour
 PT enhancers, e.g. in doughs
 PS Example 9; Page 52; 92pp; English.
 CC The present sequence represents a PCR primer for aminopeptidase
 CC (AP) II. AP is used in combination with an endopeptidase (EP) to
 CC hydrolyse polypeptides, producing protein hydrolysate (PH), useful in
 CC foods as flavour enhancer, e.g. in baked goods, enriched in: (a) Leu,
 CC Gly (especially), Glu, Ser, Asp, Asn, Pro, Cys, Ala and/or Gln, or (b),
 CC If the substrate has been deamidated, in Glu (free and/or peptide bound),
 CC in which case products are useful as animal feed additives. AP can be
 CC used in flavour-improving compositions (optionally containing EP) and in
 CC dough pre-mixes. Also AP can be used for deactivating enzymes and for
 CC converting precursors to mature proteins. AP increases the level of
 CC hydrolysis of proteins and thus of flavour development. PH have improved
 CC solubility, emulsifying and foaming properties, and products containing
 CC them have better flavour, palatability and aroma.
 SQ Sequence 23 BP; 3 A; 3 C; 9 G; 8 T;

Query Match 1.38; Score 20; DB 1; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGAGTCGCTTTGTGGGC 20
 ||||||||||||||||
 DB 4 ATGAGTCGCTTTGTGGGC 23

RESULT 7
 V82525 V82525 standard; DNA; 23 BP.
 AC V82525;
 DT 17-MAR-1999 (first entry)
 DE Aspergillus oryzae aminopeptidase II PCR forward primer #1.
 KM Aminopeptidase; protein hydrolysate; glycine releasing; protease;
 KM proteinaceous material; flavour; food; baking; animal feed additive;
 KM palatability; hydrolysis; solubility; emulsifying; foaming; aroma;
 KM mouthfeel; crust colour; PCR primer; ss.
 OS Synthetic.
 PN Aspergillus oryzae.
 PN WO9851163-A2.
 PN 15-NOV-1998.
 PF 15-MAY-1998: U09998.
 PR 16-DEC-1997: US-069719.
 PR 16-MAY-1997: US-857886.
 PR 20-OCT-1997: US-062893.
 PR 16-DEC-1997: DK-001465.
 PA (ASAHI) ASahi CHEM IND CO LTD.
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 PI Blinovsky A, Brown K, Byun T, Fujii M, Golightly E,
 PI Kofod LV, Marumotac, Mathiasen TE;
 DR WPI: 99-045177/04.
 PT Production of protein hydrolysate - using protease and enzyme that
 PT releases glycine, useful as flavour improvers in foods and animal
 PT feed additives
 PS Example 9; Page 25; 84pp; English.
 CC A method has been developed for the production of protein hydrolysates

CC (PH) comprising reacting a protein with: (i) at least one polypeptide
 CC with Gly-releasing activity; and (ii) at least one other protease so
 CC that the amount of Gly produced is greater than when the protease is
 CC used alone. PH are used to improve flavour of foods (e.g. baked goods)
 CC and as animal feed additives. PH are preferably also enriched in Glu
 CC (free and/or peptide bound), so have improved flavour and palatability.
 CC Addition of a polypeptide with Gly-releasing activity increases the
 CC degree of hydrolysis (or reduces the amount of enzyme needed) and
 CC hydrolysates have better solubility, and emulsifying and foaming
 CC properties. Baked goods containing them have improved aroma, mouthfeel
 CC and crust colour. The present sequence represents a PCR primer for
 CC Aspergillus oryzae aminopeptidase II, which is used in the method of
 CC the invention.
 SQ Sequence 23 BP; 3 A; 3 C; 9 G; 8 T;

Query Match 1.38; Score 20; DB 1; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGAGTCGCTTTGTGGGC 20
 ||||||||||||||||
 DB 4 ATGAGTCGCTTTGTGGGC 23

RESULT 8
 X00023 X00023 standard; DNA; 32 BP.
 AC X00023;
 DT 17-MAR-1999 (first entry)
 DE Aspergillus oryzae aminopeptidase II PCR forward primer #2.
 KM Aminopeptidase; protein hydrolysate; baking; dough; flavour enhancer;
 KM mouthfeel; palatability; aroma; hydrolysis; animal feed additive;
 KM PCR primer; ss.
 OS Synthetic.
 PN Aspergillus oryzae.
 PN WO9851804-A1.
 PN 19-NOV-1998.
 PF 15-MAY-1998: U09940.
 PR 20-OCT-1997: US-062893.
 PR 16-MAY-1997: US-857886.
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 PI Blinovsky A, Brown K, Byun T, Golightly E, Kofod LV;
 DR WPI: 99-045233/04.
 PT New aminopeptidase from Aspergillus oryzae - used to produce protein
 PT hydrolysates enriched in particular amino acids, used as flavour
 PT enhancers, e.g. in doughs
 PS Example 12; Page 55; 92pp; English.
 CC The present sequence represents a PCR primer for aminopeptidase
 CC (AP) II. AP is used in combination with an endopeptidase (EP) to
 CC hydrolyse polypeptides, producing protein hydrolysate (PH), useful in
 CC foods as flavour enhancer, e.g. in baked goods, enriched in: (a) Leu,
 CC Gly (especially), Glu, Ser, Asp, Asn, Pro, Cys, Ala and/or Gln, or (b),
 CC If the substrate has been deamidated, in Glu (free and/or peptide bound),
 CC in which case products are useful as animal feed additives. AP can be
 CC used in flavour-improving compositions (optionally containing EP) and in
 CC dough pre-mixes. Also AP can be used for deactivating enzymes and for
 CC converting precursors to mature proteins. AP increases the level of
 CC hydrolysis of proteins and thus of flavour development. PH have improved
 CC solubility, emulsifying and foaming properties, and products containing
 CC them have better flavour, palatability and aroma.
 SQ Sequence 32 BP; 7 A; 6 C; 8 G; 11 T;

Query Match 1.38; Score 20; DB 1; Length 32;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGAGTCGCTTTGTGGGC 20
 ||||||||||||||||
 DB 13 ATGAGTCGCTTTGTGGGC 32

```

RESULT 9
X00011
ID X00011 standard; DNA; 32 BP.
AC X00011;
DE Aspergillus oryzae aminopeptidase II PCR forward primer #2.
KW Aminopeptidase; protein hydrolysate; glycine releasing; protease;
KW proteolaccous material; flavour; food; baking; animal feed additive;
KW palatability; hydrolysis; solubility; emulsifying; foaming; aroma;
KW mouthfeel; crust colour; PCR primer; ss.
OS Synthetic.
AS Aspergillus oryzae.
PN WO9851163-A2.
PD 19-NOV-1998.
PE 15-MAY-1998; U099998.
PR 16-DEC-1997; US-069719.
PR 16-MAY-1997; US-857886.
PR 20-OCT-1997; US-062893.
PR 16-DEC-1997; DK-001465.
PA (ASAH ) ASAH CHEM IND CO LTD.
PA (NOVO ) NOVO NORDISK BIOTECH INC.
PI Blinckovsky A, Brown K, Byun T, Fujii M, Gollightly E,
PI Kofod LV, Marumotac, Mathiasen TE;
PI WPI: 99-045177/04.
PT Production of protein hydrolysate - using protease and enzyme that
PT releases glycine, useful as flavour improvers in foods and animal
PT feed additives
PS Example 12; Page 28; 84pp; English.
CC A method has been developed for the production of protein hydrolysates
CC (PH) comprising reacting a protein with: (i) at least one polypeptide
CC with gly-releasing activity; and (ii) at least one other protease so
CC that the amount of Gly produced is greater than when the protease is
CC used alone. PH are used to improve flavour of foods (e.g. baked goods)
CC and as animal feed additives. PH are preferably also enriched in Glu
CC (free and/or peptide bound), so have improved flavour and palatability.
CC Addition of a polypeptide with Gly-releasing activity increases the
CC degree of hydrolysis (or reduces the amount of enzyme needed) and
CC hydrolysates have better solubility, and emulsifying and foaming
CC properties. Baked goods containing them have improved aroma, mouthfeel
CC and crust colour. The present sequence represents a PCR primer for
CC Aspergillus oryzae aminopeptidase II, which is used in the method of
CC the invention.
SQ Sequence 32 BP; 7 A; 6 C; 8 G; 11 T;

Query Match 1.3%; Score 20; DB 1; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGTCGCTTTGTGGGC 20
    |||||||
DB 13 ATGAGTCGCTTTGTGGGC 32

RESULT 10
V55014/c
ID V55014 standard; DNA; 128 BP.
AC V55014;
DE 11-NOV-1998 (first entry)
DE Fragment of GABA gated chloride channel subunit coding sequence.
KW GABA gated chloride channel subunit; gamma amino butyric acid;
KW lepidopteran pest control; major inhibitory neurotransmitter;
KW insecticide; ss.
OS Synthetic.
OS Heliothis virescens.
PN US5767261-A.
PD 16-JUN-1998.
PE 08-NOV-1995; 554659.
PE 08-NOV-1995; US-554659.
PA (RHON ) RHONE-POULENC CO AG.
PI Wingate V, Wolff M;
PI WPI: 98-361790/31.

```

```

PT DNA encoding Heliothis GABA gated chloride channel proteins - used
PT for recombinant production of the channel for use in insecticide
PT screening
PS Example 1; Column 9; 28pp; English.
CC This sequence is a fragment of a DNA sequence of the invention. The
CC DNA sequence encodes a lepidopteran GABA (gamma amino butyric
CC acid) gated chloride channel subunit protein, specifically the Heliothis
CC virescens GABA gated chloride channel subunit protein. Cells expressing
CC the recombinant proteins, especially xenopus oocytes or Sf9 insect cells,
CC or membrane preparations of such cells, can be used to produce the
CC proteins or to screen for lepidopteran GABA-gated chloride channel
CC agonists or antagonists that may be useful as insecticides for
CC controlling lepidopteran pests, e.g. Heliothis virescens. GABA is a major
CC inhibitory neurotransmitter in mammals and insects, with the channel
CC having physiological similarity between them. However, they exhibit
CC marked differences in sensitivity to antagonists and agonists, and so the
CC recombinant channels can be used to screen for specific antagonists for
CC insects e.g. insecticides.
SQ Sequence 128 BP; 26 A; 34 C; 32 G; 36 T;

Query Match 1.3%; Score 19; DB 1; Length 128;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 TCGAGCGAGACCATGAC 323
    |||||||
DB 127 TCGAGCGAGACCATGAC 109

RESULT 11
T58840_4
Continuation (5 of 6) of T58840 from base 400001 (Mycoplasma genitalium genome.)
WP Sequence split into 6 fragments LOCUS T58840 Accession T58840
WP Fragment Name Begin End
WP T58840_0 1 110000
WP T58840_1 100001 210000
WP T58840_2 200001 310000
WP T58840_3 300001 410000
WP T58840_4 400001 510000
WP T58840_5 500001 580073

Query Match 1.3%; Score 19; DB 1; Length 110000;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 793 ATTATTAGCACTTGTC A 811
    |||||||
DB 16342 ATTATTAGCACTTGTC A 16360

RESULT 12
T91360/c
ID T91360 standard; DNA; 4372 BP.
AC T91360;
DE 14-APR-1998 (first entry)
DE Orf virus genomic DNA BamHI F and partial BamHI C fragments.
KW Parapoxvirus; vector; vaccine; antigen; HIV; herpes simplex virus;
KW Eschinococcus granulosis; Trichostrongylus; Haemonchus;
KW Osterriagta; Taenia ovis; ds.
OS Orf virus strain NZ-2.
FH key Location/Qualifiers
FH CDS complement (1..8)
FT /tag= a
FT /label= FIL
FT /note= '5' end of FIL coding sequence"
FT complement (9..49)
FT /tag= b
FT /label= FIL
FT /note= "putative FIL gene promoter"
FT complement (2421..2423)
FT /tag= c

```

FT	/label= F2L	
FT	/note= "F2L gene Initiation codon"	
FT	2437. .2451	
FT	/tag= d	
FT	/note= "intergenic region suitable for foreign	
FT	gene insertion"	
FT	complement (2424. .2463)	
FT	/tag= e	
FT	/label= PR2L	
FT	/note= "putative F2L gene promoter"	
FT	2475. .2521	
FT	/tag= f	
FT	/label= PR3R	
FT	/note= "putative F3R gene promoter"	
FT	2532. .3215	
FT	/tag= g	
FT	/label= F3R	
FT	/note= "putative F3R gene"	
FT	3183. .3237	
FT	/tag= h	
FT	/label= PR4R	
FT	/note= "putative F4R gene promoter"	
FT	3216. .3220	
FT	/tag= i	
FT	/note= "intergenic region suitable for foreign	
FT	gene insertion"	
FT	3240. .4196	
FT	/tag= j	
FT	/label= F4R	
FT	/note= "putative F4R gene"	
FT	4142. .4192	
FT	/tag= k	
FT	/label= PC1R	
FT	/note= "putative C1R gene promoter"	
FT	3573. .3597	
FT	/tag= l	
FT	/note= "intergenic region suitable for foreign	
FT	gene insertion"	
FT	4193. .4371	
FT	/tag= m	
FT	/label= C1R	
FT	/note= "putative C1R gene"	
PN	W09737031-A1.	
PD	09-OCT-1997.	
PE	27-MAR-1997; NZ0040.	
PR	29-MAR-1996; NZ-286284.	
PA	(UYOT-) UNIV OTAGO.	
PI	Lytle DJ, Robinson AJ;	
DR	WPI. 97-503116/46.	
PT	Parapoxvirus vectors containing exogenous DNA - comprise parapox	
PT	virus, especially of virus, and exogenous DNA encoding a	
PS	polypeptide, e.g. an antigen useful in vaccines	
PT	Claim 12; Fig 7; 73pp: English.	
CC	This genomic DNA molecule comprises the BamHI F fragment and part	
CC	of the BamHI C fragment of the orf virus strain NZ-2 genome	
CC	Incorporating putative promoter sequences for the F1L, F2L, F4R	
CC	and C1R genes as well as intergenic regions suitable as foreign	
CC	gene insertion sites. A claimed parapoxvirus vector comprises a	
CC	parapox virus (especially an orf virus) containing exogenous DNA.	
CC	Vector fragments or variants with equivalent immunological activity	
CC	and host cells (especially bovine or ovine testis cells)	
CC	Incorporating the vector are also claimed. The vector and	
CC	fragments/variants are useful in vaccine production (claimed),	
CC	since they can express antigens, useful in medical and veterinary	
CC	applications e.g. parasitic disease control in livestock. Particular	
CC	antigens include HIV envelope protein, herpes simplex glycoprotein,	
CC	antigens of Echinococcus granulosis, Trichostrongylus, Haemonchus,	
CC	Ostertagia and especially Tricla ovis 45W, 16 kd, 18 kd antigens.	
CC	The foreign gene is preferably incorporated in a non-essential	
CC	region of the virus genome (see F1J357-61) under control of a	
CC	poxyvirus promoter, especially of virus promoter EL1, F1L or F3R	
CC	(see F1J362-64). The orf virus vectors limit the cross-species	
CC	infection common with e.g. vaccinia virus vectors, since the virus	

```

CC has a narrow host range: orf is also less virulent than vaccinia in
CC man.
SQ Sequence 4372 BP; 817 A; 1346 C; 1439 G; 770 T;

Query Match 1.2%; Score 18; DB 1; Length 4372;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 332 CCAGCGTCGAGTCACCG 349
|||||
DB 862 CCAGCGTCGAGTCACCG 845

RESULT 13
X40591
ID ID X40591 standard; cDNA; 444 BP.
AC X40591:
DT 18-JUN-1999 (first entry)
DE Human secreted protein 5' EST SEQ ID No: 191.
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide; prostate;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition; ds.
OS Homo sapiens.
PN M09906550-A2.
PD 11-FEB-1999.
PE 31-JUL-1998; IB1232.
PF 01-AUG-1997; US-905144.
PA (GEST ) GENSET.
PI Duclert A, Dumas Mline Edwards J, Lacroix B;
DR WPJ; 99-153780/13.
DR P-PSDB; Y11869.
PT New isolated prostate-derived nucleic acids - used to develop
PT products which may have cytokine, immune regulatory, haematopoiesis
PT regulatory, anti-inflammatory or tumour inhibition activity
PS Claim 1; Page 300; 675pp; English.
PS X40438 to X40715 represent 5' expressed sequence tags (ESTs) for human
CC secreted proteins expressed in prostate, and encode the proteins given in
CC Y11716 to Y11993 respectively. The proteins given represent the signal
CC peptide and an N-terminal fragment of a secreted protein. The nucleic
CC acid sequences can be used for producing secreted human gene products.
CC They can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell proliferation and
CC differentiation activity, haematopoiesis regulating activity, tissue
CC growth regulating activity, reproductive hormone regulating activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, tumour inhibition
CC activity or other activities. The products can be used in forensic, gene
CC therapy and chromosome mapping procedures. The sequences can also be used
CC for obtaining corresponding promoter sequences. The nucleic acids
CC encoding the signal peptides can be used for directing extracellular
CC secretion of a polypeptide or the insertion of a polypeptide into a
CC membrane, or importing a polypeptide into a cell.
SQ Sequence 444 BP; 109 A; 116 C; 132 G; 83 T;

Query Match 1.1%; Score 17; DB 1; Length 444;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 203 ACGAGAGCTGAGAGAG 219
|||||
DB 284 ACGAGAGCTGAGAGAG 300

RESULT 14
T95755
ID ID T95755 standard; cDNA; 1231 BP.
AC T95755;
DT 11-MAY-1998 (first entry)

```

```

DE Arabidopsis SCARECROW clone SRP3.
KW SCARECROW; SCR: SRP3 gene; transgenic plant; root; gravitropism;
OS Arabidopsis thaliana.
FH Key Location/Qualifiers
FT CDS 1..978
FT     /*tag= a
PN M09741152-A1.
PD 06-NOV-1997.
PF 25-APR-1997; U07022.
PR 24-APR-1997; US-842445.
PR 26-APR-1996; US-638617.
PA (UNY) UNIV NEW YORK STATE.
PI Befrey PN, Dilaurenzio L, Helariutta Y, Malamy JE,
PI Pysch L, Mysiocka-Diller J;
DR WPI: 97-549683/50.
DR P-PSDB: W8180.
PT DNA encoding Arabidopsis SCARECROW protein - useful to modify plant
PT cell division and therefore alter root development, or alter plant
PT stem or hypocotyl gravitropism
PS Claim 6; Fig 9; 22pp; English.
CC This sequence comprises a partial cDNA clone of the Arabidopsis
CC SCARECROW (SCR) SRP3 gene, identified by homology to the
CC Arabidopsis SCR gene (see T95753). It encodes a 325-amino acid
CC polypeptide (see W8180). SCR is expressed specifically in embryo
CC root progenitor tissue and in certain root and stem tissues. It
CC regulates a specific asymmetric division, and controls gravitropic
CC response in aerial structures and root formation. Nucleic acid
CC molecules (see T95753-66) encoding SCR protein sequences (see
CC W8178-201), SCR proteins lacking 1 to 4 of MOTIFs I to VI, or
CC SCR MOTIF I, II, III, IV, V or VI are claimed. Transgenic plants
CC can be engineered to overexpress SCR, so that cell division is
CC increased in roots, resulting in thicker root development, while a
CC plant with an altered stem or hypocotyl gravitropism is less
CC susceptible to lodging. Plants that contain an antisense
CC molecule that suppresses the expression of endogenous SCR gene
CC product show thinner root development. A gene of interest can be
CC placed under control of a SCR promoter and expressed in a plant to
CC confer herbicide, salt, pathogen or insect resistance, or when
CC expressed in stems to increase starch, lignin or cellulose
CC biosynthesis (all claimed).
SQ Sequence 1231 BP; 323 A; 228 C; 318 G; 362 T;

Query Match 1.1%; Score 17; DB 1; Length 1231;
Best Local Similarity 100.0%; Pred. NO. 48;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 578 TCGGATCAGCTTGAG 594
    ||||||||||||||||
DB 949 TCGGATCAGCTTGAG 965

RESULT 15
002845
ID 002845 standard; cDNA; 1507 BP.
AC 002845.
DT 01-JUN-1990 (first entry)
DE Complete sequence of a cDNA encoding porcine brain natriuretic peptide.
KW Porcine BNP; natriuretic; diuretic; vasodilator; hypertension; ds.
OS Sus scrofa.
FH Key Location/Qualifiers
FT exon 100..222
FT     /*tag= a
FT intron 223..468
FT     /*tag= b
FT exon 469..723
FT     /*tag= c
FT intron 724..1275
FT     /*tag= d
FT exon 1276..1289
FT     /*tag= e
FT mat_protein 661..723

```

```

FT FT mat_protein /*tag= f
FT 1276..1289
FT     /*tag= g
PN W08912069-A.
PD 14-DEC-1989.
PF 31-MAY-1989; 02373.
PR 31-MAY-1988; US-200383.
PR 14-JUN-1988; US-2006470.
PR 19-JAN-1989; US-299880.
PA (CALB) Calif Biotech Inc.
PI Sellhammer JJ, Lewicki J, Scarborough RM, Porter GU;
PI WPI:90-007453/01.
DR P-PSDB: R04084.
PT New natriuretic and vasodilator peptides - obtd. using cDNA sequence
PT encoding porcine brain natriuretic peptide and related human canine genes
PS Disclosure; P; English.
CC Sequences derived from porcine brain natriuretic gene can be used in part
CC as probes to isolate similar genes from different species. Although human
CC natriuretic related peptides (NRP) could not be isolated directly, canine
CC NRPs could be, and these could then be used as probes to their human
CC equivalents.
CC See also Q02851, Q02846, Q02849 and Q02850.
SQ Sequence 1507 BP; 322 A; 450 C; 407 G; 328 T;

```

```

Query Match 1.1%; Score 17; DB 1; Length 1507;
Best Local Similarity 100.0%; Pred. NO. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 GTTGCTGGAGGCGC 55
    ||||||||||||||||
DB 251 GTTGCTGGAGGCGC 267

```

Search completed: May 11, 2000, 00:37:54
Job time: 7776 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2000, 22:26:26 ; Search time 62.31 Seconds

(Without alignments)

3038.293 Million cpi1 updates/sec

Title: US-09-080-127-1

Perfect score: 1491

Sequence: 1 ATGACGTGCGCTTTTGTGGGC.....ATTCTCACTGAGG-ATTAG 1491

Scoring table: OLIGO-MTC

Gapop 60.0 , Gapext 60.0

Searched: 226296 seqs, 63486255 residues

Word size : 0

Total number of hits satisfying chosen parameters: 452592

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database :

Issued Patents, NA:*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/5C_COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/5D_COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/6_COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/PCMTUS_COMB.seq:*
- 7: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	19	1.3	128	2	US-08-554-659-7
C 2	17	1.1	1488	1	US-08-140-104A-4
C 3	17	1.1	1497	1	US-08-140-104A-3
4	17	1.1	1504	4	US-08-850-910A-17
5	17	1.1	1507	4	US-08-850-910A-38
C 6	17	1.1	1509	1	US-08-140-104A-1
C 7	17	1.1	5089	7	5177197-31
C 8	16	1.1	225	1	US-08-030-731A-40
C 9	16	1.1	285	3	US-08-716-942-23
C 10	16	1.1	729	6	PCT-US96-05320A-1376
C 11	16	1.1	1524	7	5512669-1
C 12	16	1.1	1525	7	5229279-1
C 13	16	1.1	1734	7	5352575-8
C 14	16	1.1	1842	1	US-08-442-134A-1
C 15	16	1.1	1842	1	US-08-444-581B-1
C 16	16	1.1	1842	2	US-08-446-088A-1
C 17	16	1.1	1997	4	US-08-987-466-2
C 18	16	1.1	2084	2	US-08-720-899-3
C 19	16	1.1	2084	2	US-08-459-610-3
C 20	16	1.1	2084	3	US-08-343-804-3
C 21	16	1.1	2084	4	US-08-687-399-3
C 22	16	1.1	2084	4	US-08-600-908A-3
C 23	16	1.1	2084	5	US-08-683-838A-3
C 24	16	1.1	2256	3	US-08-896-590A-3
C 25	16	1.1	2259	3	US-08-896-590A-1
C 26	16	1.1	2550	4	US-08-658-665-42

27	16	1.1	2550	4	US-08-658-665-44	Sequence 44, Appl
28	16	1.1	2724	1	US-08-349-006-1	Sequence 1, Appl
29	16	1.1	2724	4	US-08-658-665-37	Sequence 37, Appl
30	16	1.1	2724	6	PCT-US94-04180-1	Sequence 1, Appl
C 31	16	1.1	2972	4	US-08-720-484A-3	Sequence 3, Appl
C 32	16	1.1	3425	3	US-08-768-147B-1	Sequence 1, Appl
C 33	16	1.1	3627	1	US-08-104-072B-6	Sequence 6, Appl
C 34	16	1.1	3627	2	US-08-351-413-7	Sequence 7, Appl
C 35	16	1.1	3627	4	US-09-025-583-7	Sequence 7, Appl
C 36	16	1.1	4260	4	US-08-658-665-38	Sequence 38, Appl
C 37	16	1.1	4594	4	US-08-658-665-43	Sequence 43, Appl
C 38	16	1.1	4594	4	US-08-658-665-45	Sequence 45, Appl
C 39	16	1.1	4766	4	US-08-658-665-41	Sequence 41, Appl
C 40	16	1.1	4875	2	US-08-460-739-1	Sequence 1, Appl
C 41	16	1.1	5234	4	US-08-638-665-13	Sequence 73, Appl
C 42	16	1.1	7091	4	US-08-658-665-40	Sequence 40, Appl
C 43	16	1.1	10095	5	US-08-822-586-45	Sequence 45, Appl
C 44	16	1.1	44377	3	US-08-804-227C-7	Sequence 7, Appl
C 45	16	1.1	44377	4	US-08-804-198-1	Sequence 1, Appl
C 46	15	1.0	33	3	US-08-432-871C-26	Sequence 26, Appl
C 47	15	1.0	284	1	US-08-150-203A-2	Sequence 2, Appl
C 48	15	1.0	284	2	US-08-454-730-2	Sequence 2, Appl
C 49	15	1.0	312	4	US-08-853-659A-30	Sequence 30, Appl
C 50	15	1.0	515	4	US-08-975-316-42	Sequence 22, Appl
C 51	15	1.0	519	4	US-08-853-659A-21	Sequence 21, Appl
C 52	15	1.0	527	1	US-07-998-973A-22	Sequence 22, Appl
C 53	15	1.0	527	4	US-08-452-800-22	Sequence 22, Appl
C 54	15	1.0	527	6	PCT-US92-11353-22	Sequence 22, Appl
C 55	15	1.0	606	3	US-08-432-871C-48	Sequence 48, Appl
C 56	15	1.0	624	3	US-08-713-000-9	Sequence 9, Appl
C 57	15	1.0	624	4	US-08-975-316-9	Sequence 9, Appl
C 58	15	1.0	676	3	US-08-533-669A-3	Sequence 3, Appl
C 59	15	1.0	684	4	US-08-975-316-45	Sequence 45, Appl
C 60	15	1.0	737	3	US-08-861-549-2	Sequence 2, Appl
C 61	15	1.0	1067	2	US-08-264-534-7	Sequence 7, Appl
C 62	15	1.0	1067	2	US-08-465-500-7	Sequence 7, Appl
C 63	15	1.0	1067	3	US-08-346-126-7	Sequence 7, Appl
C 64	15	1.0	1067	3	US-08-346-128-7	Sequence 7, Appl
C 65	15	1.0	1161	6	PCT-US93-11153-31	Sequence 31, Appl
C 66	15	1.0	1176	6	US-08-387-942C-17	Sequence 17, Appl
C 67	15	1.0	1289	3	US-08-344-833-1	Sequence 1, Appl
C 68	15	1.0	1371	3	US-08-871-924A-3	Sequence 3, Appl
C 69	15	1.0	1371	4	US-08-643-034A-3	Sequence 3, Appl
C 70	15	1.0	1488	6	PCT-US92-06965A-14	Sequence 14, Appl
C 71	15	1.0	1573	1	US-08-597-545-4	Sequence 4, Appl
C 72	15	1.0	1573	1	US-08-457-135-4	Sequence 4, Appl
C 73	15	1.0	1573	4	US-08-966-316-10	Sequence 10, Appl
C 74	15	1.0	1637	4	US-08-528-199-5	Sequence 2, Appl
C 75	15	1.0	1704	2	US-08-528-199-5	Sequence 2, Appl
C 76	15	1.0	1704	2	US-08-313-181-2	Sequence 2, Appl
C 77	15	1.0	1894	5	US-08-749-816-1	Sequence 1, Appl
C 78	15	1.0	2012	6	PCT-US94-10621-2	Sequence 2, Appl
C 79	15	1.0	2012	6	PCT-US94-02233-2	Sequence 2, Appl
C 80	15	1.0	2104	2	US-08-592-126-96	Sequence 96, Appl
C 81	15	1.0	2104	2	US-08-633-879C-17	Sequence 17, Appl
C 82	15	1.0	2121	4	US-08-633-879C-3	Sequence 3, Appl
C 83	15	1.0	2194	4	US-07-618-946B-24	Sequence 24, Appl
C 84	15	1.0	2217	2	US-08-712-241-4	Sequence 4, Appl
C 85	15	1.0	2217	2	US-08-153-848-27	Sequence 27, Appl
C 86	15	1.0	2254	6	PCT-US93-11153-27	Sequence 27, Appl
C 87	15	1.0	2310	2	US-08-417-330A-15	Sequence 15, Appl
C 88	15	1.0	2348	4	US-08-876-546A-1	Sequence 1, Appl
C 89	15	1.0	2442	4	US-07-618-946B-25	Sequence 25, Appl
C 90	15	1.0	2556	2	US-08-725-012-1	Sequence 1, Appl
C 91	15	1.0	2620	1	US-08-056-200-109	Sequence 109, App
C 92	15	1.0	2620	4	US-08-800-644-109	Sequence 109, App
C 93	15	1.0	2620	4	US-08-867-941-2	Sequence 2, Appl
C 94	15	1.0	2722	5	US-08-804-439A-10	Sequence 10, Appl
C 95	15	1.0	2722	5	US-08-720-229-10	Sequence 10, Appl
C 96	15	1.0	2729	1	US-07-938-782A-1	Sequence 1, Appl
C 97	15	1.0	2729	2	US-08-630-524-1	Sequence 1, Appl
C 98	15	1.0	2729	2	PCT-US93-08131-1	Sequence 1, Appl
C 99	15	1.0	2729	6		Sequence 1, Appl

100	15	1.0	3435	2	US-08-366-577-1	Sequence 1, Appl1	c 173	14	0.9	78	3	US-08-790-963-88	Sequence 88, Appl1
101	15	1.0	3435	6	PCT-US96-00005-1	Sequence 1, Appl1	174	14	0.9	110	1	US-08-258-026a-26	Sequence 26, Appl1
102	15	1.0	3760	4	US-08-724-354D-1	Sequence 1, Appl1	175	14	0.9	117	1	US-08-136-993-3	Sequence 7, Appl1
c 103	15	1.0	4054	1	US-07-618-946B-2	Sequence 2, Appl1	176	14	0.9	138	3	US-08-611-757-17	Sequence 17, Appl1
c 104	15	1.0	4245	3	US-08-929-967-4	Sequence 4, Appl1	177	14	0.9	138	6	PCT-US95-05980-17	Sequence 17, Appl1
c 105	15	1.0	4411	3	US-08-929-967-5	Sequence 5, Appl1	c 178	14	0.9	148	1	US-08-136-993-22	Sequence 22, Appl1
c 106	15	1.0	5134	4	US-08-635-121-1	Sequence 1, Appl1	179	14	0.9	152	1	US-08-136-993-20	Sequence 20, Appl1
c 107	15	1.0	5339	2	US-08-064-121-1	Sequence 1, Appl1	180	14	0.9	158	6	PCT-US95-02576-7	Sequence 7, Appl1
c 108	15	1.0	5339	5	US-08-478-015-1	Sequence 1, Appl1	c 181	14	0.9	164	1	US-07-766-351-16	Sequence 16, Appl1
c 109	15	1.0	5339	2	US-08-475-975-1	Sequence 1, Appl1	c 182	14	0.9	164	1	US-08-059-032-1	Sequence 16, Appl1
c 110	15	1.0	5838	4	US-08-850-880-3	Sequence 3, Appl1	c 183	14	0.9	164	6	PCT-US91-07290-16	Sequence 16, Appl1
c 111	15	1.0	5838	4	US-08-944-916-3	Sequence 3, Appl1	184	14	0.9	169	1	US-08-320-559-19	Sequence 19, Appl1
c 112	15	1.0	5838	4	US-08-814-877-3	Sequence 3, Appl1	185	14	0.9	169	1	US-08-327-392-19	Sequence 19, Appl1
c 113	15	1.0	5889	2	US-08-463-092B-5	Sequence 5, Appl1	186	14	0.9	189	6	PCT-US94-04496-19	Sequence 19, Appl1
c 114	15	1.0	5889	3	US-08-462-109A-5	Sequence 5, Appl1	187	14	0.9	269	1	US-08-025-038-31	Sequence 31, Appl1
c 115	15	1.0	5889	3	US-08-460-907B-5	Sequence 5, Appl1	188	14	0.9	269	1	US-08-025-038-32	Sequence 32, Appl1
c 116	15	1.0	5889	5	US-08-463-179A-5	Sequence 5, Appl1	189	14	0.9	269	1	US-08-025-038-33	Sequence 33, Appl1
c 117	15	1.0	5889	5	US-08-461-384B-5	Sequence 5, Appl1	190	14	0.9	269	1	US-08-025-038-34	Sequence 34, Appl1
c 118	15	1.0	6353	3	US-08-929-967-6	Sequence 6, Appl1	191	14	0.9	269	1	US-08-025-038-35	Sequence 35, Appl1
c 119	15	1.0	6449	3	US-08-479-537A-4	Sequence 4, Appl1	192	14	0.9	269	1	US-08-025-038-36	Sequence 36, Appl1
c 120	15	1.0	6555	2	US-08-351-413-2	Sequence 2, Appl1	193	14	0.9	269	1	US-08-050-073-19	Sequence 19, Appl1
c 121	15	1.0	6555	4	US-09-025-583-2	Sequence 2, Appl1	194	14	0.9	269	1	US-08-050-073-20	Sequence 20, Appl1
c 122	15	1.0	7399	3	US-08-418-848A-9	Sequence 9, Appl1	195	14	0.9	269	1	US-08-050-073-21	Sequence 21, Appl1
123	15	1.0	7650	4	US-08-867-941-1	Sequence 1, Appl1	196	14	0.9	269	1	US-08-050-073-22	Sequence 22, Appl1
124	15	1.0	7650	4	US-08-188-583-5	Sequence 5, Appl1	197	14	0.9	269	1	US-08-050-073-39	Sequence 39, Appl1
125	15	1.0	9709	5	US-08-388-353-1	Sequence 1, Appl1	198	14	0.9	269	2	US-08-116-942-7	Sequence 7, Appl1
126	15	1.0	9709	5	US-08-488-551B-1	Sequence 1, Appl1	c 199	14	0.9	282	2	US-08-450-945-62	Sequence 62, Appl1
127	15	1.0	9739	1	US-08-022-835-1	Sequence 1, Appl1	200	14	0.9	285	3	US-08-716-942-6	Sequence 6, Appl1
128	15	1.0	9739	1	US-08-388-809-1	Sequence 1, Appl1	201	14	0.9	325	3	US-08-320-559-23	Sequence 23, Appl1
129	15	1.0	9739	3	US-08-647-714-1	Sequence 1, Appl1	202	14	0.9	336	1	US-08-327-392-23	Sequence 23, Appl1
130	15	1.0	9746	1	US-08-022-835-3	Sequence 3, Appl1	203	14	0.9	336	6	PCT-US94-04496-23	Sequence 23, Appl1
131	15	1.0	9746	1	US-08-388-809-3	Sequence 3, Appl1	204	14	0.9	412	1	US-08-318-193-88	Sequence 88, Appl1
132	15	1.0	9746	3	US-08-647-714-3	Sequence 3, Appl1	205	14	0.9	433	1	US-07-971-101-5	Sequence 5, Appl1
c 133	15	1.0	9960	5	US-08-822-586-46	Sequence 46, Appl1	c 206	14	0.9	440	1	US-08-076-726-12	Sequence 12, Appl1
c 134	15	1.0	10095	5	US-08-822-586-45	Sequence 45, Appl1	c 207	14	0.9	440	1	US-08-076-726-13	Sequence 13, Appl1
c 135	15	1.0	12588	4	US-08-387-942C-1	Sequence 1, Appl1	c 208	14	0.9	450	1	US-08-485-971-8	Sequence 8, Appl1
136	15	1.0	14985	2	US-08-652-972a-6	Sequence 6, Appl1	c 209	14	0.9	450	1	US-08-485-971-9	Sequence 9, Appl1
137	15	1.0	14985	6	PCT-US96-06231A-6	Sequence 6, Appl1	c 210	14	0.9	450	1	US-08-260-452-5	Sequence 5, Appl1
c 138	15	1.0	15512	4	US-08-853-659A-5	Sequence 5, Appl1	c 211	14	0.9	450	1	US-08-260-452-6	Sequence 6, Appl1
c 139	15	1.0	15512	4	US-08-853-659A-6	Sequence 6, Appl1	c 212	14	0.9	450	1	US-08-260-452-7	Sequence 7, Appl1
c 140	15	1.0	15512	4	US-08-853-659A-3	Sequence 3, Appl1	c 213	14	0.9	450	1	US-08-275-876-8	Sequence 8, Appl1
141	15	1.0	15512	4	US-08-853-659A-66	Sequence 66, Appl1	c 214	14	0.9	450	1	US-08-275-876-9	Sequence 9, Appl1
142	15	1.0	15512	5	US-08-846-538-35	Sequence 35, Appl1	c 215	14	0.9	450	2	US-08-383-754-8	Sequence 8, Appl1
143	15	1.0	24417	4	US-08-846-538-1	Sequence 1, Appl1	c 216	14	0.9	450	2	US-08-383-754-9	Sequence 9, Appl1
c 144	15	1.0	24701	4	US-08-853-659A-2	Sequence 2, Appl1	c 217	14	0.9	450	2	US-08-485-978-8	Sequence 8, Appl1
c 145	15	1.0	24701	4	US-08-853-659A-3	Sequence 3, Appl1	c 218	14	0.9	450	2	US-08-485-978-9	Sequence 9, Appl1
c 146	15	1.0	24701	4	US-08-853-659A-60	Sequence 60, Appl1	c 219	14	0.9	450	3	US-08-481-970-5	Sequence 5, Appl1
147	15	1.0	24701	4	US-08-853-659A-61	Sequence 61, Appl1	c 220	14	0.9	450	3	US-08-481-970-6	Sequence 6, Appl1
c 148	15	1.0	28804	3	US-08-592-874-1	Sequence 1, Appl1	c 221	14	0.9	450	3	US-08-486-814-8	Sequence 8, Appl1
c 149	15	1.0	28804	5	US-09-096-942-2	Sequence 2, Appl1	c 222	14	0.9	450	4	US-08-486-814-9	Sequence 9, Appl1
c 150	15	1.0	30001	1	US-08-125-468-1	Sequence 1, Appl1	c 223	14	0.9	450	4	US-08-487-472-8	Sequence 8, Appl1
c 151	15	1.0	30001	3	US-08-474-933-1	Sequence 1, Appl1	c 224	14	0.9	450	4	US-08-487-472-9	Sequence 9, Appl1
c 152	15	1.0	35100	2	US-08-306-691B-19	Sequence 19, Appl1	c 225	14	0.9	450	4	US-08-897-719-5	Sequence 5, Appl1
c 153	15	1.0	35100	6	PCT-US93-06251-19	Sequence 19, Appl1	c 226	14	0.9	450	4	US-08-897-719-6	Sequence 6, Appl1
c 154	15	1.0	246240	3	US-08-724-394A-20	Sequence 20, Appl1	c 227	14	0.9	450	6	PCT-US95-08179-9	Sequence 9, Appl1
c 155	15	1.0	246240	3	US-08-724-394A-21	Sequence 21, Appl1	c 228	14	0.9	481	2	US-08-694-652-1	Sequence 2, Appl1
c 156	15	1.0	246240	3	US-08-724-394A-22	Sequence 22, Appl1	c 229	14	0.9	486	3	US-08-199-984-7	Sequence 7, Appl1
157	14	0.9	18	2	US-08-358-995-17	Sequence 17, Appl1	c 230	14	0.9	495	2	US-08-249-671A-6	Sequence 6, Appl1
c 158	14	0.9	20	5	US-08-294-424-23	Sequence 23, Appl1	c 231	14	0.9	500	4	US-08-487-472-7	Sequence 7, Appl1
c 159	14	0.9	20	5	US-08-841-267-2	Sequence 2, Appl1	c 232	14	0.9	500	6	PCT-US95-08179-7	Sequence 7, Appl1
c 160	14	0.9	27	1	US-08-366-783-21	Sequence 21, Appl1	c 233	14	0.9	500	1	US-08-485-971-7	Sequence 7, Appl1
161	14	0.9	27	1	US-08-846-021A-27	Sequence 27, Appl1	c 234	14	0.9	520	2	US-08-383-754-7	Sequence 7, Appl1
162	14	0.9	29	2	US-08-601-435-4	Sequence 4, Appl1	c 235	14	0.9	520	2	US-08-485-978-7	Sequence 7, Appl1
163	14	0.9	29	4	US-08-931-047-4	Sequence 4, Appl1	c 236	14	0.9	520	4	US-08-486-814-7	Sequence 7, Appl1
164	14	0.9	29	4	US-08-783-302-4	Sequence 4, Appl1	c 237	14	0.9	520	4	US-08-487-472-7	Sequence 7, Appl1
c 165	14	0.9	33	1	US-08-138-608-25	Sequence 25, Appl1	c 238	14	0.9	520	6	PCT-US95-08179-7	Sequence 7, Appl1
c 166	14	0.9	36	1	US-08-435-350-3	Sequence 3, Appl1	c 239	14	0.9	540	7	5310729-1	Patent No. 5310729
c 167	14	0.9	40	5	US-08-841-267-4	Sequence 4, Appl1	c 241	14	0.9	569	1	US-08-485-971-6	Sequence 6, Appl1
c 168	14	0.9	69	3	US-08-790-963-80	Sequence 80, Appl1	c 242	14	0.9	569	1	US-08-275-876-6	Sequence 6, Appl1
c 169	14	0.9	69	3	US-08-790-963-81	Sequence 81, Appl1	c 243	14	0.9	569	1	US-08-275-876-6	Sequence 6, Appl1
c 170	14	0.9	76	1	US-07-753-110B-12	Sequence 12, Appl1	c 244	14	0.9	569	2	US-08-383-754-6	Sequence 6, Appl1
c 171	14	0.9	76	1	US-08-503-730-16	Sequence 16, Appl1	c 245	14	0.9	569	2	US-08-383-754-6	Sequence 6, Appl1
c 172	14	0.9	76	3	US-08-507-634-13	Sequence 13, Appl1							

246	14	0.9	569	2	US-08-485-978-6	Sequence 6, Appli	C 319	14	0.9	1161	4	US-09-006-491-5	Sequence 5, Appli
C 247	14	0.9	569	2	US-08-485-978-6	Sequence 6, Appli	C 320	14	0.9	1176	6	PCT-US95-03032-2	Sequence 2, Appli
C 248	14	0.9	569	3	US-08-486-814-6	Sequence 6, Appli	C 321	14	0.9	1225	1	US-08-181-629A-3	Sequence 3, Appli
C 249	14	0.9	569	3	US-08-486-814-6	Sequence 6, Appli	C 322	14	0.9	1225	1	US-08-286-020-1	Sequence 1, Appli
C 250	14	0.9	569	4	US-08-487-472-6	Sequence 6, Appli	C 323	14	0.9	1225	2	US-08-603-919-1	Sequence 1, Appli
C 251	14	0.9	569	4	US-08-487-472-6	Sequence 6, Appli	C 324	14	0.9	1229	1	US-08-304-076A-1	Sequence 1, Appli
C 252	14	0.9	569	6	PCT-US95-08179-6	Sequence 6, Appli	C 325	14	0.9	1245	4	US-08-945-296-1	Sequence 1, Appli
C 253	14	0.9	569	6	PCT-US95-08179-6	Sequence 6, Appli	C 326	14	0.9	1254	1	US-08-236-918A-1	Sequence 1, Appli
C 254	14	0.9	573	3	US-08-290-665A-118	Sequence 118, App	C 327	14	0.9	1257	1	US-07-946-232-7	Sequence 7, Appli
C 255	14	0.9	573	3	US-08-290-665A-123	Sequence 123, App	C 328	14	0.9	1272	1	US-08-073-799C-7	Sequence 7, Appli
C 256	14	0.9	573	6	PCT-US95-10398-118	Sequence 118, App	C 329	14	0.9	1272	1	US-07-947-672-7	Sequence 7, Appli
C 257	14	0.9	573	6	PCT-US95-10398-123	Sequence 123, App	C 330	14	0.9	1272	3	US-08-432-043-7	Sequence 7, Appli
C 258	14	0.9	600	1	US-07-753-738B-11	Sequence 11, Appli	C 331	14	0.9	1272	3	US-08-660-963-7	Sequence 7, Appli
C 259	14	0.9	606	3	US-08-416-603-9	Sequence 9, Appli	C 332	14	0.9	1278	4	US-08-765-268A-2	Sequence 2, Appli
C 260	14	0.9	616	1	US-08-029-022-1	Sequence 1, Appli	C 333	14	0.9	1289	2	US-08-644-664B-13	Sequence 13, Appli
C 261	14	0.9	616	1	US-08-029-022-3	Sequence 1, Appli	C 334	14	0.9	1289	4	US-08-761-277-13	Sequence 13, Appli
C 262	14	0.9	616	2	US-08-246-376-1	Sequence 1, Appli	C 335	14	0.9	1298	2	US-08-473-157A-2	Sequence 2, Appli
C 263	14	0.9	616	2	US-08-246-376-3	Sequence 3, Appli	C 336	14	0.9	1302	2	US-08-719-331-7	Sequence 7, Appli
C 264	14	0.9	616	3	US-07-972-135-1	Sequence 1, Appli	C 337	14	0.9	1306	3	US-08-827-190-12	Sequence 12, Appli
C 265	14	0.9	616	3	US-07-972-135-3	Sequence 3, Appli	C 338	14	0.9	1320	2	US-08-599-252-84	Sequence 84, Appli
C 266	14	0.9	616	5	US-08-256-004-1	Sequence 1, Appli	C 339	14	0.9	1320	2	US-08-436-074-57	Sequence 57, Appli
C 267	14	0.9	616	5	US-08-256-004-3	Sequence 3, Appli	C 340	14	0.9	1320	3	PCT-US96-06352-84	Sequence 84, Appli
C 268	14	0.9	616	6	PCT-US93-05366-1	Sequence 1, Appli	C 341	14	0.9	1320	6	PCT-US96-06352-84	Sequence 84, Appli
C 269	14	0.9	616	6	PCT-US93-05366-3	Sequence 3, Appli	C 342	14	0.9	1320	6	PCT-US96-06583-84	Sequence 84, Appli
C 270	14	0.9	627	1	US-07-753-738B-1	Sequence 1, Appli	C 343	14	0.9	1386	1	US-08-247-475-16	Sequence 16, Appli
C 271	14	0.9	627	1	US-07-753-738B-4	Sequence 4, Appli	C 344	14	0.9	1386	1	US-08-247-475-16	Sequence 16, Appli
C 272	14	0.9	627	1	US-07-753-738B-7	Sequence 7, Appli	C 345	14	0.9	1386	1	US-08-479-650-16	Sequence 16, Appli
C 273	14	0.9	627	1	US-07-753-738B-8	Sequence 8, Appli	C 346	14	0.9	1386	1	US-08-479-650-16	Sequence 16, Appli
C 274	14	0.9	627	1	US-07-753-738B-9	Sequence 9, Appli	C 347	14	0.9	1386	2	US-08-191-866D-20	Sequence 20, Appli
C 275	14	0.9	627	1	US-07-753-738B-10	Sequence 10, Appli	C 348	14	0.9	1386	2	US-08-191-866D-20	Sequence 20, Appli
C 276	14	0.9	687	2	US-08-028-463-6	Sequence 6, Appli	C 349	14	0.9	1386	2	US-08-674-169-16	Sequence 16, Appli
C 277	14	0.9	687	2	US-08-461-836-6	Sequence 6, Appli	C 350	14	0.9	1386	2	US-08-674-169-16	Sequence 16, Appli
C 278	14	0.9	720	7	5310729-3	Patent No. 5310729	C 351	14	0.9	1386	3	US-08-185-949B-20	Sequence 20, Appli
C 279	14	0.9	739	1	US-07-872-678A-2	Sequence 2, Appli	C 352	14	0.9	1386	3	US-08-185-949B-20	Sequence 20, Appli
C 280	14	0.9	757	1	US-08-229-515A-10	Sequence 10, Appli	C 353	14	0.9	1428	1	US-08-051-142-1	Sequence 1, Appli
C 281	14	0.9	757	1	US-08-645-865-10	Sequence 10, Appli	C 354	14	0.9	1428	1	US-08-624-601-7	Sequence 7, Appli
C 282	14	0.9	765	1	US-08-366-783-3	Sequence 3, Appli	C 355	14	0.9	1448	1	US-08-585-758A-1	Sequence 1, Appli
C 283	14	0.9	771	6	PCT-US95-12987-1	Sequence 1, Appli	C 356	14	0.9	1448	2	US-08-977-818-1	Sequence 1, Appli
C 284	14	0.9	771	6	PCT-US95-12987-3	Sequence 3, Appli	C 357	14	0.9	1448	2	US-08-670-274B-9	Sequence 9, Appli
C 285	14	0.9	771	6	PCT-US95-12987-5	Sequence 5, Appli	C 358	14	0.9	1472	1	US-08-123-161A-9	Sequence 9, Appli
C 286	14	0.9	793	5	US-08-971-158-2	Sequence 2, Appli	C 359	14	0.9	1472	1	US-08-483-278-9	Sequence 9, Appli
C 287	14	0.9	825	3	US-08-781-550-42	Sequence 42, Appli	C 360	14	0.9	1495	2	US-08-471-791-19	Sequence 19, Appli
C 288	14	0.9	826	7	5252477-2	Patent No. 5252477	C 361	14	0.9	1495	6	PCT-US91-01746-19	Sequence 19, Appli
C 289	14	0.9	827	3	US-08-781-550-41	Sequence 41, Appli	C 362	14	0.9	1497	3	US-08-820-170A-41	Sequence 41, Appli
C 290	14	0.9	840	2	US-08-595-559-4	Sequence 4, Appli	C 363	14	0.9	1528	3	US-08-945-296-3	Sequence 3, Appli
C 291	14	0.9	869	2	US-08-249-671A-7	Sequence 7, Appli	C 364	14	0.9	1541	2	US-08-200-232-3	Sequence 3, Appli
C 292	14	0.9	877	4	US-08-989-289-2	Sequence 2, Appli	C 365	14	0.9	1541	6	PCT-US95-02219-3	Sequence 3, Appli
C 293	14	0.9	879	2	US-08-249-671A-10	Sequence 10, Appli	C 366	14	0.9	1541	6	PCT-US95-02219A-3	Sequence 3, Appli
C 294	14	0.9	915	2	US-08-556-806-5	Sequence 5, Appli	C 367	14	0.9	1558	2	US-08-467-607-2	Sequence 2, Appli
C 295	14	0.9	930	1	US-08-029-022-2	Sequence 2, Appli	C 368	14	0.9	1558	3	US-08-469-362-2	Sequence 2, Appli
C 296	14	0.9	930	1	US-08-029-022-4	Sequence 4, Appli	C 369	14	0.9	1558	3	US-08-850-392-2	Sequence 2, Appli
C 297	14	0.9	930	2	US-08-246-376-2	Sequence 2, Appli	C 370	14	0.9	1566	4	US-08-875-802-1	Sequence 1, Appli
C 298	14	0.9	930	2	US-08-246-376-4	Sequence 4, Appli	C 371	14	0.9	1574	5	US-09-173-581-12	Sequence 12, Appli
C 299	14	0.9	930	3	US-07-972-135-2	Sequence 2, Appli	C 372	14	0.9	1607	4	US-08-749-902-4	Sequence 4, Appli
C 300	14	0.9	930	3	US-07-972-135-4	Sequence 4, Appli	C 373	14	0.9	1615	1	US-08-176-413-4	Sequence 4, Appli
C 301	14	0.9	930	5	US-08-256-004-2	Sequence 2, Appli	C 374	14	0.9	1615	6	PCT-US94-14919-4	Sequence 4, Appli
C 302	14	0.9	930	6	PCT-US93-05366-2	Sequence 2, Appli	C 375	14	0.9	1624	2	US-08-416-870C-5	Sequence 5, Appli
C 303	14	0.9	931	6	PCT-US93-05366-4	Sequence 4, Appli	C 376	14	0.9	1633	1	US-08-097-827-1	Sequence 1, Appli
C 304	14	0.9	931	1	US-08-203-905B-4	Sequence 4, Appli	C 377	14	0.9	1633	2	US-08-494-574-1	Sequence 1, Appli
C 305	14	0.9	948	1	US-08-589-446-5	Sequence 5, Appli	C 378	14	0.9	1645	4	US-08-461-812-1	Sequence 1, Appli
C 306	14	0.9	948	1	US-08-444-882-5	Sequence 5, Appli	C 379	14	0.9	1650	4	US-08-776-210-6	Sequence 6, Appli
C 307	14	0.9	948	2	US-08-389-459A-5	Sequence 5, Appli	C 380	14	0.9	1655	4	US-08-945-296-2	Sequence 2, Appli
C 308	14	0.9	1014	6	PCT-US92-10421-1	Sequence 1, Appli	C 381	14	0.9	1665	1	US-08-440-377A-1	Sequence 1, Appli
C 309	14	0.9	1035	2	US-08-891-254-8	Sequence 8, Appli	C 382	14	0.9	1665	2	US-08-687-852-1	Sequence 1, Appli
C 310	14	0.9	1035	3	US-08-819-539-8	Sequence 8, Appli	C 383	14	0.9	1680	6	PCT-US93-06748-1	Sequence 1, Appli
C 311	14	0.9	1035	4	US-09-030-270A-8	Sequence 8, Appli	C 384	14	0.9	1702	1	US-07-616-022C-1	Sequence 1, Appli
C 312	14	0.9	1035	6	PCT-US96-08819-8	Sequence 8, Appli	C 385	14	0.9	1704	1	US-08-097-282B-1	Sequence 1, Appli
C 313	14	0.9	1051	4	US-08-865-273-1	Sequence 1, Appli	C 386	14	0.9	1746	2	US-08-235-471-8	Sequence 8, Appli
C 314	14	0.9	1051	4	US-08-487-135B-17	Sequence 17, Appli	C 387	14	0.9	1756	4	US-08-465-640-1	Sequence 1, Appli
C 315	14	0.9	1054	4	US-08-915-972A-17	Sequence 17, Appli	C 388	14	0.9	1769	4	US-08-765-268A-1	Sequence 3, Appli
C 316	14	0.9	1054	4	US-09-177-909-17	Sequence 17, Appli	C 389	14	0.9	1785	4	US-08-559-505-3	Sequence 3, Appli
C 317	14	0.9	1060	3	US-08-385-335A-14	Sequence 14, Appli	C 390	14	0.9	1785	4	US-08-749-907-3	Sequence 3, Appli
C 318	14	0.9	1161	2	US-08-663-310-5	Sequence 5, Appli	C 391	14	0.9	1785	6	PCT-US95-07721-7	Sequence 7, Appli

C 392	14	0.9	1789	3	US-08-366-490-1	Sequence 1, Appl1	C 465	14	0.9	2225	2	US-08-450-945-56	Sequence 56, Appl1
C 393	14	0.9	1794	6	PCT-US94-10080-9	Sequence 9, Appl1	C 466	14	0.9	2232	6	PCT-US96-05320A-113	Sequence 113, Appl1
C 394	14	0.9	1825	4	US-08-176-414B-1	Sequence 1, Appl1	C 467	14	0.9	2265	3	US-08-940-332-1	Sequence 1, Appl1
C 395	14	0.9	1825	4	US-08-336-724-1	Sequence 1, Appl1	C 468	14	0.9	2304	1	US-08-464-266-1	Sequence 1, Appl1
C 396	14	0.9	1827	4	US-08-824-878-2	Sequence 2, Appl1	C 469	14	0.9	2304	1	US-08-464-266-1	Sequence 1, Appl1
C 397	14	0.9	1833	4	US-08-365-470-2	Sequence 2, Appl1	C 470	14	0.9	2304	1	US-08-464-266-1	Sequence 1, Appl1
C 398	14	0.9	1844	1	US-07-872-644-26	Sequence 26, Appl1	C 471	14	0.9	2343	4	US-09-018-760-3	Sequence 3, Appl1
C 399	14	0.9	1844	1	US-08-297-494-26	Sequence 26, Appl1	C 472	14	0.9	2389	4	US-08-372-652-8	Sequence 8, Appl1
C 400	14	0.9	1844	1	US-08-297-510-26	Sequence 26, Appl1	C 473	14	0.9	2389	6	PCT-US95-16311-8	Sequence 8, Appl1
C 401	14	0.9	1844	2	US-08-479-532-26	Sequence 26, Appl1	C 474	14	0.9	2391	4	US-08-817-900-1	Sequence 1, Appl1
C 402	14	0.9	1844	2	US-08-455-526-26	Sequence 26, Appl1	C 475	14	0.9	2464	1	US-07-863-169A-4	Sequence 4, Appl1
C 403	14	0.9	1844	2	US-08-455-526-26	Sequence 26, Appl1	C 476	14	0.9	2464	1	US-08-429-964-4	Sequence 4, Appl1
C 404	14	0.9	1844	5	US-09-139-491-26	Sequence 26, Appl1	C 477	14	0.9	2464	6	PCT-US93-08062-4	Sequence 4, Appl1
C 405	14	0.9	1844	6	PCT-US92-03222-26	Sequence 26, Appl1	C 478	14	0.9	2521	1	US-08-188-228-59	Sequence 59, Appl1
C 406	14	0.9	1848	1	US-08-447-422-15	Sequence 15, Appl1	C 479	14	0.9	2521	1	US-08-332-643-53	Sequence 53, Appl1
C 407	14	0.9	1864	4	US-08-468-812-3	Sequence 3, Appl1	C 480	14	0.9	2521	1	US-08-332-643-53	Sequence 53, Appl1
C 408	14	0.9	1876	4	US-08-483-332-24	Sequence 24, Appl1	C 481	14	0.9	2564	5	US-08-276-968A-19	Sequence 19, Appl1
C 409	14	0.9	1876	3	US-08-485-538A-24	Sequence 24, Appl1	C 482	14	0.9	2582	4	US-08-816-105A-2	Sequence 4, Appl1
C 410	14	0.9	1876	4	US-08-910-041-24	Sequence 24, Appl1	C 483	14	0.9	2605	3	US-08-680-395-4	Sequence 4, Appl1
C 411	14	0.9	1882	1	US-08-458-120-1	Sequence 1, Appl1	C 484	14	0.9	2624	1	US-08-032-382B-1	Sequence 1, Appl1
C 412	14	0.9	1882	4	US-08-867-970-1	Sequence 1, Appl1	C 485	14	0.9	2739	3	US-07-743-357-16	Sequence 16, Appl1
C 413	14	0.9	1883	1	US-08-202-056-2	Sequence 2, Appl1	C 486	14	0.9	2793	3	US-07-646-537B-1	Sequence 1, Appl1
C 414	14	0.9	1899	2	US-08-884-072-4	Sequence 4, Appl1	C 487	14	0.9	2793	3	US-08-795-868-13	Sequence 13, Appl1
C 415	14	0.9	1917	2	US-08-458-477A-1	Sequence 1, Appl1	C 488	14	0.9	2812	2	US-08-920-812-16	Sequence 16, Appl1
C 416	14	0.9	1917	4	US-09-033-153-1	Sequence 1, Appl1	C 489	14	0.9	2812	2	US-08-920-812-16	Sequence 16, Appl1
C 417	14	0.9	1933	1	US-08-076-093A-1	Sequence 1, Appl1	C 490	14	0.9	2812	2	US-08-921-177-16	Sequence 16, Appl1
C 418	14	0.9	1933	1	US-08-410-451-1	Sequence 1, Appl1	C 491	14	0.9	2812	2	US-08-362-577C-16	Sequence 16, Appl1
C 419	14	0.9	1933	1	US-08-410-455-1	Sequence 1, Appl1	C 492	14	0.9	2812	2	US-08-920-828-16	Sequence 16, Appl1
C 420	14	0.9	1933	1	US-08-418-919-1	Sequence 1, Appl1	C 493	14	0.9	2866	2	US-08-663-310-7	Sequence 7, Appl1
C 421	14	0.9	1933	2	US-08-410-453A-2	Sequence 2, Appl1	C 494	14	0.9	2866	2	US-09-006-491-7	Sequence 7, Appl1
C 422	14	0.9	1933	2	US-08-701-265-1	Sequence 1, Appl1	C 495	14	0.9	3033	3	US-07-743-357-19	Sequence 19, Appl1
C 423	14	0.9	1933	2	US-08-410-454A-2	Sequence 2, Appl1	C 496	14	0.9	3063	3	US-08-142-439A-1	Sequence 1, Appl1
C 424	14	0.9	1933	3	US-08-284-586-1	Sequence 1, Appl1	C 497	14	0.9	3066	3	US-08-869-477-1	Sequence 1, Appl1
C 425	14	0.9	1933	3	US-08-410-456A-2	Sequence 2, Appl1	C 498	14	0.9	3125	3	US-08-037-816A-13	Sequence 13, Appl1
C 426	14	0.9	1933	3	US-08-805-478-1	Sequence 1, Appl1	C 499	14	0.9	3125	3	US-08-530-146-13	Sequence 13, Appl1
C 427	14	0.9	1933	3	US-08-802-627A-1	Sequence 1, Appl1	C 500	14	0.9	3147	4	US-08-781-802-7	Sequence 7, Appl1
C 428	14	0.9	1933	4	US-08-801-238-1	Sequence 1, Appl1	C 501	14	0.9	3182	2	US-08-595-559-6	Sequence 6, Appl1
C 429	14	0.9	1933	4	US-08-801-228-1	Sequence 1, Appl1	C 502	14	0.9	3252	4	US-08-809-740A-1	Sequence 1, Appl1
C 430	14	0.9	1933	6	PCT-US94-06380-1	Sequence 1, Appl1	C 503	14	0.9	3252	4	US-08-809-740A-1	Sequence 1, Appl1
C 431	14	0.9	1951	2	US-08-910-973-16	Sequence 16, Appl1	C 504	14	0.9	3264	3	US-08-708-541A-31	Sequence 31, Appl1
C 432	14	0.9	1970	2	US-08-028-463-14	Sequence 14, Appl1	C 505	14	0.9	3264	3	US-08-708-541A-33	Sequence 33, Appl1
C 433	14	0.9	1970	2	US-08-461-836-14	Sequence 14, Appl1	C 506	14	0.9	3377	7	5198345-16	Sequence 33, Appl1
C 434	14	0.9	1983	1	US-08-073-799C-9	Sequence 9, Appl1	C 507	14	0.9	3402	2	US-08-480-917-1	Sequence 1, Appl1
C 435	14	0.9	2038	3	US-08-736-770-2	Sequence 2, Appl1	C 508	14	0.9	3456	2	US-08-190-687B-24	Sequence 24, Appl1
C 436	14	0.9	2048	1	US-07-602-608-11	Sequence 11, Appl1	C 509	14	0.9	3533	3	US-08-476-062A-40	Sequence 40, Appl1
C 437	14	0.9	2048	1	US-08-261-578-11	Sequence 11, Appl1	C 510	14	0.9	3533	6	PCT-US96-01314-40	Sequence 40, Appl1
C 438	14	0.9	2057	3	US-08-820-170A-42	Sequence 42, Appl1	C 511	14	0.9	3533	7	5424399-1	Sequence 40, Appl1
C 439	14	0.9	2070	1	US-07-766-351-6	Sequence 6, Appl1	C 512	14	0.9	3561	2	US-08-097-997A-12	Sequence 12, Appl1
C 440	14	0.9	2070	6	PCT-US91-07290-6	Sequence 6, Appl1	C 513	14	0.9	3581	3	US-08-738-349-1	Sequence 1, Appl1
C 441	14	0.9	2085	1	US-08-059-032-6	Sequence 1, Appl1	C 514	14	0.9	3656	1	US-08-393-734-1	Sequence 1, Appl1
C 442	14	0.9	2085	1	US-08-072-070-1	Sequence 1, Appl1	C 515	14	0.9	3812	1	US-08-188-582-19	Sequence 19, Appl1
C 443	14	0.9	2085	1	US-08-465-746-1	Sequence 1, Appl1	C 516	14	0.9	3812	1	US-08-646-715-19	Sequence 19, Appl1
C 444	14	0.9	2085	2	US-08-214-164-1	Sequence 1, Appl1	C 517	14	0.9	3840	1	US-08-462-092-1	Sequence 1, Appl1
C 445	14	0.9	2085	2	US-08-469-434-1	Sequence 1, Appl1	C 518	14	0.9	3840	5	US-08-746-822-1	Sequence 1, Appl1
C 446	14	0.9	2085	2	US-08-214-422-1	Sequence 1, Appl1	C 519	14	0.9	3840	7	5510474-1	Sequence 1, Appl1
C 447	14	0.9	2085	2	US-08-467-852A-1	Sequence 1, Appl1	C 520	14	0.9	3854	1	US-08-365-470-1	Sequence 1, Appl1
C 448	14	0.9	2085	3	US-08-468-718-1	Sequence 1, Appl1	C 521	14	0.9	3856	3	US-07-743-357-20	Sequence 20, Appl1
C 449	14	0.9	2085	4	US-08-246-636-1	Sequence 1, Appl1	C 522	14	0.9	3858	3	US-08-344-155C-98	Sequence 98, Appl1
C 450	14	0.9	2085	4	US-08-247-636-1	Sequence 1, Appl1	C 523	14	0.9	3863	7	5217870-1	Sequence 8, Appl1
C 451	14	0.9	2085	4	US-08-319-795-1	Sequence 1, Appl1	C 524	14	0.9	3901	1	US-08-326-297-1	Sequence 21, Appl1
C 452	14	0.9	2085	4	US-08-468-985-1	Sequence 1, Appl1	C 525	14	0.9	3901	5	US-08-617-454-1	Sequence 1, Appl1
C 453	14	0.9	2086	5	US-08-312-949-1	Sequence 1, Appl1	C 526	14	0.9	3901	6	PCT-US94-01144-1	Sequence 1, Appl1
C 454	14	0.9	2095	1	US-08-333-901-5	Sequence 5, Appl1	C 527	14	0.9	3958	3	US-07-952-853-21	Sequence 21, Appl1
C 455	14	0.9	2095	2	US-08-405-230-8	Sequence 8, Appl1	C 528	14	0.9	3958	4	US-08-914-848-21	Sequence 21, Appl1
C 456	14	0.9	2095	2	US-08-456-582-5	Sequence 5, Appl1	C 529	14	0.9	4080	2	US-08-570-311-7	Sequence 7, Appl1
C 457	14	0.9	2095	3	US-08-910-590-8	Sequence 8, Appl1	C 530	14	0.9	4080	3	US-08-353-485-7	Sequence 7, Appl1
C 458	14	0.9	2095	3	US-08-898-789-5	Sequence 5, Appl1	C 531	14	0.9	4094	4	US-08-841-483-5	Sequence 5, Appl1
C 459	14	0.9	2139	3	US-08-611-280-3	Sequence 3, Appl1	C 532	14	0.9	4113	4	US-07-743-357-21	Sequence 21, Appl1
C 460	14	0.9	2158	1	US-07-602-608-1	Sequence 1, Appl1	C 533	14	0.9	4134	3	US-08-817-090B-3	Sequence 1, Appl1
C 461	14	0.9	2158	1	US-08-261-578-1	Sequence 1, Appl1	C 534	14	0.9	4137	3	US-08-817-090B-3	Sequence 1, Appl1
C 462	14	0.9	2179	3	US-08-551-356-3	Sequence 3, Appl1	C 535	14	0.9	4154	1	US-08-131-365B-37	Sequence 37, Appl1
C 463	14	0.9	2179	6	PCT-US93-12687-3	Sequence 3, Appl1	C 536	14	0.9	4154	3	US-08-668-123-37	Sequence 37, Appl1
C 464	14	0.9	2214	1	US-07-985-458-1	Sequence 1, Appl1	C 537	14	0.9	4283	1	US-08-343-401A-3	Sequence 3, Appl1

C 538	14	0.9	4283	2	US-08-445-265A-1	Sequence 1, Appl1	C 611	14	0.9	6226	2	US-08-542-363-1	Sequence 1, Appl1
539	14	0.9	4307	2	US-08-190-687B-7	Sequence 7, Appl1	C 612	14	0.9	6238	5	US-08-646-58-5	Sequence 5, Appl1
540	14	0.9	4348	4	US-08-915-868-1	Sequence 1, Appl1	C 613	14	0.9	6244	1	US-08-076-728-15	Sequence 15, Appl1
541	14	0.9	4394	2	US-08-095-734-1	Sequence 1, Appl1	C 614	14	0.9	6244	1	US-08-260-452-8	Sequence 8, Appl1
542	14	0.9	4394	3	US-08-444-623-1	Sequence 1, Appl1	C 615	14	0.9	6244	4	US-08-481-970-8	Sequence 8, Appl1
543	14	0.9	4394	3	US-08-471-869-1	Sequence 1, Appl1	C 616	14	0.9	6244	4	US-08-897-719-8	Sequence 8, Appl1
544	14	0.9	4394	6	PCT-US94-08267-1	Sequence 1, Appl1	C 617	14	0.9	6253	5	US-08-893-327-15	Sequence 15, Appl1
C 545	14	0.9	4455	3	US-08-474-169-5	Sequence 5, Appl1	C 618	14	0.9	6280	5	US-08-893-327-17	Sequence 17, Appl1
C 546	14	0.9	4473	4	US-09-048-804-1	Sequence 1, Appl1	C 619	14	0.9	6280	5	US-08-893-327-19	Sequence 19, Appl1
C 547	14	0.9	4503	4	US-08-770-301A-2	Sequence 2, Appl1	C 620	14	0.9	6295	4	US-08-659-206A-4	Sequence 4, Appl1
C 548	14	0.9	4530	1	US-08-229-515A-9	Sequence 9, Appl1	C 621	14	0.9	6315	3	US-08-808-793-2	Sequence 2, Appl1
C 549	14	0.9	4530	1	US-08-645-865-9	Sequence 9, Appl1	C 622	14	0.9	6315	3	US-08-808-793-2	Sequence 2, Appl1
C 550	14	0.9	4534	4	US-08-935-450-1	Sequence 1, Appl1	C 623	14	0.9	6315	5	US-08-772-512A-2	Sequence 2, Appl1
C 551	14	0.9	4566	3	US-08-465-976A-1	Sequence 1, Appl1	C 624	14	0.9	6315	5	US-08-772-512A-2	Sequence 2, Appl1
C 552	14	0.9	4566	4	US-08-982-412-1	Sequence 1, Appl1	C 625	14	0.9	6318	5	US-08-808-793-1	Sequence 1, Appl1
553	14	0.9	4586	5	US-09-031-563-26	Sequence 26, Appl1	C 626	14	0.9	6318	5	US-08-772-512A-1	Sequence 1, Appl1
554	14	0.9	4623	1	US-08-370-193A-7	Sequence 7, Appl1	C 627	14	0.9	6387	1	US-07-721-775A-1	Sequence 1, Appl1
C 555	14	0.9	4696	3	US-08-929-967-15	Sequence 15, Appl1	C 628	14	0.9	6387	1	US-08-339-658-1	Sequence 1, Appl1
C 556	14	0.9	4732	1	US-07-884-811-1	Sequence 1, Appl1	C 629	14	0.9	6395	4	US-08-687-559-2	Sequence 2, Appl1
C 557	14	0.9	4732	1	US-07-885-971-1	Sequence 1, Appl1	C 630	14	0.9	6395	4	US-08-338-702-7	Sequence 7, Appl1
C 558	14	0.9	4732	1	US-08-087-783A-1	Sequence 1, Appl1	C 631	14	0.9	6513	1	US-08-337-339-7	Sequence 7, Appl1
C 559	14	0.9	4732	1	US-08-194-088B-1	Sequence 1, Appl1	C 632	14	0.9	6513	1	US-08-151-574-31	Sequence 31, Appl1
C 560	14	0.9	4732	1	US-08-194-088B-1	Sequence 1, Appl1	C 633	14	0.9	6513	6	PCT-US95-14282-7	Sequence 7, Appl1
C 561	14	0.9	4732	6	PCT-US93-04648-1	Sequence 1, Appl1	C 634	14	0.9	6513	6	PCT-US95-14378-7	Sequence 7, Appl1
562	14	0.9	4887	1	US-07-789-915A-7	Sequence 7, Appl1	C 635	14	0.9	6519	3	US-08-808-793-24	Sequence 24, Appl1
563	14	0.9	4887	1	US-08-005-002C-7	Sequence 7, Appl1	C 636	14	0.9	6660	4	US-08-771-602D-1	Sequence 1, Appl1
564	14	0.9	4887	1	US-08-487-203A-7	Sequence 7, Appl1	C 637	14	0.9	6718	4	US-08-962-284-1	Sequence 1, Appl1
C 565	14	0.9	4928	1	US-08-345-913-1	Sequence 1, Appl1	C 638	14	0.9	6756	1	US-08-419-448-31	Sequence 31, Appl1
C 566	14	0.9	4937	1	US-08-038-682-3	Sequence 3, Appl1	C 639	14	0.9	6756	3	US-08-419-448-31	Sequence 31, Appl1
C 567	14	0.9	4937	1	US-08-302-832-3	Sequence 3, Appl1	C 640	14	0.9	6763	3	US-08-756-506-23	Sequence 23, Appl1
C 568	14	0.9	4937	3	US-08-530-198-3	Sequence 3, Appl1	C 641	14	0.9	6827	6	US-08-222-616-17	Sequence 17, Appl1
C 569	14	0.9	4937	3	US-08-468-880-3	Sequence 3, Appl1	C 642	14	0.9	6827	6	PCT-US95-04228-17	Sequence 17, Appl1
C 570	14	0.9	4937	4	US-08-728-470-3	Sequence 3, Appl1	C 643	14	0.9	6854	2	US-08-468-036-4	Sequence 4, Appl1
C 571	14	0.9	4937	4	US-08-617-697-3	Sequence 3, Appl1	C 644	14	0.9	6854	3	US-08-376-843-4	Sequence 4, Appl1
C 572	14	0.9	4963	1	US-08-076-726-16	Sequence 16, Appl1	C 645	14	0.9	6889	1	US-08-286-740-1	Sequence 2, Appl1
C 573	14	0.9	4963	1	US-08-260-452-9	Sequence 9, Appl1	C 646	14	0.9	6889	6	PCT-US95-09576-2	Sequence 2, Appl1
C 574	14	0.9	4963	3	US-08-481-970-9	Sequence 9, Appl1	C 647	14	0.9	7287	1	US-08-659-206A-1	Sequence 1, Appl1
C 575	14	0.9	4963	4	US-08-897-719-9	Sequence 9, Appl1	C 648	14	0.9	7305	1	US-08-286-740-4	Sequence 4, Appl1
C 576	14	0.9	5030	3	US-08-588-201-14	Sequence 14, Appl1	C 649	14	0.9	7305	6	PCT-US95-09576-4	Sequence 4, Appl1
C 577	14	0.9	5030	4	US-09-169-605-14	Sequence 14, Appl1	C 650	14	0.9	7326	2	US-08-194-466-1	Sequence 1, Appl1
C 578	14	0.9	5060	4	US-08-893-327-12	Sequence 12, Appl1	C 651	14	0.9	7360	1	US-08-286-740-1	Sequence 1, Appl1
C 579	14	0.9	5060	4	US-08-656-555-2	Sequence 2, Appl1	C 652	14	0.9	7360	6	PCT-US95-09576-1	Sequence 1, Appl1
580	14	0.9	5062	4	US-08-656-555-1	Sequence 1, Appl1	C 653	14	0.9	7607	1	US-08-222-616-19	Sequence 19, Appl1
C 581	14	0.9	5157	3	US-08-474-169-7	Sequence 7, Appl1	C 654	14	0.9	7607	1	PCT-US93-04228-19	Sequence 19, Appl1
C 582	14	0.9	5158	3	US-08-929-967-16	Sequence 16, Appl1	C 655	14	0.9	7852	5	US-08-836-022A-2	Sequence 2, Appl1
583	14	0.9	5178	3	US-08-474-169-2	Sequence 2, Appl1	C 656	14	0.9	7897	5	US-07-916-096A-10	Sequence 10, Appl1
584	14	0.9	5187	3	US-08-540-406-3	Sequence 3, Appl1	C 657	14	0.9	7897	5	US-08-836-022A-1	Sequence 1, Appl1
585	14	0.9	5187	3	US-08-540-406-9	Sequence 9, Appl1	C 658	14	0.9	8236	2	US-08-461-837-1	Sequence 1, Appl1
586	14	0.9	5187	5	US-08-656-055-3	Sequence 3, Appl1	C 659	14	0.9	8236	5	US-08-973-222-1	Sequence 1, Appl1
587	14	0.9	5187	5	US-08-656-055-9	Sequence 9, Appl1	C 660	14	0.9	8236	6	PCT-US96-09485-1	Sequence 1, Appl1
588	14	0.9	5187	6	PCT-US95-13233-3	Sequence 3, Appl1	C 661	14	0.9	8299	2	US-08-462-014-2	Sequence 2, Appl1
589	14	0.9	5187	6	PCT-US95-13233-9	Sequence 9, Appl1	C 662	14	0.9	8378	6	PCT-US91-09055-1	Sequence 1, Appl1
C 590	14	0.9	5288	5	US-08-540-406-18	Sequence 18, Appl1	C 663	14	0.9	8509	2	US-08-462-014-1	Sequence 1, Appl1
C 591	14	0.9	5288	5	US-08-656-055-18	Sequence 18, Appl1	C 664	14	0.9	8533	1	US-07-846-181-6	Sequence 6, Appl1
C 592	14	0.9	5288	6	PCT-US95-13233-18	Sequence 18, Appl1	C 665	14	0.9	8533	1	US-07-846-181-6	Sequence 6, Appl1
593	14	0.9	5362	5	US-08-463-210-5	Sequence 5, Appl1	C 666	14	0.9	8573	6	US-07-845-989-6	Sequence 6, Appl1
594	14	0.9	5461	5	US-07-998-289B-7	Sequence 7, Appl1	C 667	14	0.9	8573	6	PCT-US92-08228-6	Sequence 6, Appl1
C 595	14	0.9	5496	1	US-08-181-629A-2	Sequence 2, Appl1	C 668	14	0.9	9108	6	PCT-US95-04228-45	Sequence 45, Appl1
C 596	14	0.9	5585	3	US-08-305-221-1	Sequence 1, Appl1	C 669	14	0.9	9207	5	US-08-388-353-800	Sequence 800, App
C 597	14	0.9	5620	4	US-08-793-170-21	Sequence 21, Appl1	C 670	14	0.9	9207	5	US-08-488-551B-800	Sequence 800, App
C 598	14	0.9	5676	3	US-08-663-998-3	Sequence 3, Appl1	C 671	14	0.9	9323	1	US-08-038-688-6	Sequence 6, Appl1
C 599	14	0.9	5682	3	US-08-663-998-4	Sequence 4, Appl1	C 672	14	0.9	9323	1	US-08-302-832-6	Sequence 6, Appl1
C 600	14	0.9	5900	3	US-08-663-998-1	Sequence 1, Appl1	C 673	14	0.9	9323	3	US-08-530-198-6	Sequence 6, Appl1
C 601	14	0.9	5943	1	US-08-206-176-1	Sequence 1, Appl1	C 674	14	0.9	9323	3	US-08-469-880-6	Sequence 6, Appl1
C 602	14	0.9	5952	3	US-08-663-998-2	Sequence 2, Appl1	C 675	14	0.9	9323	4	US-08-728-470-6	Sequence 6, Appl1
C 603	14	0.9	6002	2	US-08-698-551-15	Sequence 15, Appl1	C 676	14	0.9	9323	4	US-08-617-697-6	Sequence 6, Appl1
604	14	0.9	6002	3	US-08-602-228-15	Sequence 15, Appl1	C 677	14	0.9	9370	1	US-08-320-555-27	Sequence 27, Appl1
605	14	0.9	6002	3	US-08-839-032A-15	Sequence 15, Appl1	C 678	14	0.9	9370	6	PCT-US94-04486-27	Sequence 27, Appl1
606	14	0.9	6063	1	US-08-193-744-4	Sequence 4, Appl1	C 679	14	0.9	9391	1	US-08-320-555-25	Sequence 25, Appl1
607	14	0.9	6063	3	US-08-788-279-4	Sequence 4, Appl1	C 680	14	0.9	10398	6	PCT-US94-04486-25	Sequence 25, Appl1
C 608	14	0.9	6180	2	US-08-386-727-1	Sequence 1, Appl1	C 681	14	0.9	10398	3	US-08-331-384-1	Sequence 1, Appl1
C 609	14	0.9	6180	4	US-08-600-452A-1	Sequence 1, Appl1	C 682	14	0.9	10398	3	US-08-708-188-1	Sequence 1, Appl1
610	14	0.9	6206	3	US-08-474-169-3	Sequence 3, Appl1	C 683	14	0.9	10398	3	US-08-836-087-1	Sequence 1, Appl1

C 664	14	0.9	10596	1	US-07-884-811-15	Sequence 15, Appl	C 757	13	0.9	20	2	US-08-470-354-4	Sequence 4, Appl
C 665	14	0.9	10596	1	US-07-885-971-15	Sequence 15, Appl	C 758	13	0.9	20	2	US-08-479-383-4	Sequence 4, Appl
C 666	14	0.9	10596	1	US-08-087-783A-15	Sequence 15, Appl	C 759	13	0.9	20	3	US-08-479-041-4	Sequence 4, Appl
C 667	14	0.9	10596	1	US-08-194-088B-15	Sequence 15, Appl	C 760	13	0.9	20	3	US-08-708-541A-18	Sequence 18, Appl
C 668	14	0.9	10596	3	US-08-119-087-15	Sequence 15, Appl	C 761	13	0.9	20	6	PCT-US96-00108-5	Sequence 5, Appl
C 669	14	0.9	10596	6	PCT-US93-04648-15	Sequence 15, Appl	C 762	13	0.9	22	2	US-08-592-126-16	Sequence 16, Appl
C 670	14	0.9	10627	1	US-08-060-925A-12	Sequence 12, Appl	C 763	13	0.9	22	3	US-08-117-952-132	Sequence 432, App
C 671	14	0.9	10807	1	US-08-206-176-7	Sequence 7, Appl	C 764	13	0.9	22	4	US-08-687-080-16	Sequence 16, Appl
C 672	14	0.9	11517	2	US-07-920-281C-1	Sequence 5, Appl	C 765	13	0.9	23	2	US-08-171-718-72	Sequence 72, Appl
C 673	14	0.9	11832	3	US-08-416-603-1	Sequence 1, Appl	C 766	13	0.9	25	2	US-08-579-667-11	Sequence 11, Appl
C 674	14	0.9	11832	3	US-08-276-852-156	Sequence 156, App	C 767	13	0.9	25	2	US-08-426-792-22	Sequence 22, Appl
C 675	14	0.9	11254	1	US-08-276-852-170	Sequence 170, App	C 768	13	0.9	25	4	US-08-260-546-5	Sequence 5, Appl
C 676	14	0.9	11254	2	US-08-899-575-156	Sequence 156, App	C 769	13	0.9	25	6	PCT-US96-01643-10	Sequence 10, Appl
C 677	14	0.9	13254	2	US-08-899-575-170	Sequence 170, App	C 770	13	0.9	27	2	US-08-425-299A-11	Sequence 11, Appl
C 678	14	0.9	13254	2	US-08-899-575-156	Sequence 156, App	C 771	13	0.9	27	2	US-08-758-306-694	Sequence 694, App
C 679	14	0.9	13254	2	US-08-899-575-170	Sequence 170, App	C 772	13	0.9	27	3	US-08-465-095-3	Sequence 3, Appl
C 701	14	0.9	13254	6	PCT-US95-08743-156	Sequence 156, App	C 773	13	0.9	27	6	PCT-US94-00300-3	Sequence 3, Appl
C 702	14	0.9	11254	6	PCT-US95-08743-170	Sequence 170, App	C 774	13	0.9	29	2	US-08-596-024-1	Sequence 1, Appl
C 703	14	0.9	11254	6	US-08-646-538-35	Sequence 35, Appl	C 775	13	0.9	31	4	US-08-935-886-6	Sequence 6, Appl
C 704	14	0.9	15664	1	US-08-402-282-3	Sequence 3, Appl	C 776	13	0.9	31	4	US-08-858-623A-23	Sequence 23, Appl
C 705	14	0.9	15664	1	US-08-508-004-3	Sequence 3, Appl	C 777	13	0.9	31	6	PCT-US94-03796-14	Sequence 14, Appl
C 706	14	0.9	15664	1	US-08-402-066-3	Sequence 3, Appl	C 778	13	0.9	31	7	5464743-2	Patent No. 5464743
C 707	14	0.9	15664	1	US-08-402-066-3	Sequence 3, Appl	C 779	13	0.9	33	4	US-08-858-623A-12	Sequence 12, Appl
C 708	14	0.9	15664	1	US-08-402-068-3	Sequence 3, Appl	C 780	13	0.9	34	1	US-08-075-193-12	Sequence 12, Appl
C 709	14	0.9	15894	1	US-08-348-891A-1	Sequence 1, Appl	C 781	13	0.9	34	1	US-08-071-881-86	Sequence 86, Appl
C 710	14	0.9	16075	5	US-09-096-942-1	Sequence 1, Appl	C 782	13	0.9	34	2	US-08-741-881-89	Sequence 89, Appl
C 711	14	0.9	16885	2	US-08-390-878-16	Sequence 16, Appl	C 783	13	0.9	34	2	US-08-739-158-86	Sequence 86, Appl
C 712	14	0.9	17410	1	US-07-841-646-3	Sequence 3, Appl	C 784	13	0.9	34	2	US-08-739-158-89	Sequence 89, Appl
C 713	14	0.9	17410	1	US-08-147-023-3	Sequence 3, Appl	C 785	13	0.9	34	3	US-08-739-167-86	Sequence 86, Appl
C 714	14	0.9	17410	2	US-08-447-570-3	Sequence 3, Appl	C 786	13	0.9	34	3	US-08-739-167-89	Sequence 89, Appl
C 715	14	0.9	17410	3	US-08-449-700-3	Sequence 3, Appl	C 787	13	0.9	34	4	US-08-564-090A-12	Sequence 12, Appl
C 716	14	0.9	17410	6	PCT-US95-07349-1	Sequence 3, Appl	C 788	13	0.9	34	5	US-08-404-796-86	Sequence 86, Appl
C 717	14	0.9	17415	6	US-08-449-699A-3	Sequence 3, Appl	C 789	13	0.9	34	5	US-08-404-796-89	Sequence 89, Appl
C 718	14	0.9	19182	4	US-08-850-880-11	Sequence 11, Appl	C 790	13	0.9	34	5	US-08-931-869-86	Sequence 86, Appl
C 719	14	0.9	19182	4	US-08-944-916-11	Sequence 11, Appl	C 791	13	0.9	34	5	US-08-931-869-89	Sequence 89, Appl
C 720	14	0.9	19307	5	US-08-836-022A-10	Sequence 10, Appl	C 792	13	0.9	34	6	PCT-US94-06698-12	Sequence 12, Appl
C 721	14	0.9	24979	6	US-08-147-777-3	Sequence 3, Appl	C 793	13	0.9	37	1	US-08-131-365B-52	Sequence 52, Appl
C 722	14	0.9	24979	6	PCT-US93-03985-3	Sequence 3, Appl	C 794	13	0.9	37	2	US-08-468-700-13	Sequence 13, Appl
C 723	14	0.9	28958	1	US-08-258-261B-6	Sequence 6, Appl	C 795	13	0.9	37	3	US-08-668-123-52	Sequence 52, Appl
C 724	14	0.9	28958	1	US-08-456-837-6	Sequence 6, Appl	C 796	13	0.9	37	4	US-08-704-706A-14	Sequence 14, Appl
C 725	14	0.9	28958	1	US-08-457-312-6	Sequence 6, Appl	C 797	13	0.9	38	1	US-08-360-841-9	Sequence 9, Appl
C 726	14	0.9	28958	1	US-08-457-646A-6	Sequence 6, Appl	C 798	13	0.9	39	1	US-08-360-841-6	Sequence 6, Appl
C 727	14	0.9	28958	2	US-08-458-076A-6	Sequence 6, Appl	C 799	13	0.9	39	3	US-08-533-298-7	Sequence 7, Appl
C 728	14	0.9	28958	2	US-08-764-233A-4	Sequence 4, Appl	C 800	13	0.9	39	3	US-08-533-298-9	Sequence 9, Appl
C 729	14	0.9	28958	2	US-08-457-335A-6	Sequence 6, Appl	C 801	13	0.9	40	2	US-08-753-054-17	Sequence 17, Appl
C 730	14	0.9	28958	2	US-08-729-214-6	Sequence 6, Appl	C 802	13	0.9	40	4	US-08-850-049-126	Sequence 126, App
C 731	14	0.9	34303	3	US-08-735-609-4	Sequence 4, Appl	C 803	13	0.9	40	4	US-08-050-478-126	Sequence 126, App
C 732	14	0.9	34382	3	US-08-374-483-6	Sequence 6, Appl	C 804	13	0.9	45	1	US-08-044-506B-14	Sequence 14, Appl
C 733	14	0.9	35935	4	US-08-735-609-1	Sequence 1, Appl	C 806	13	0.9	45	2	US-08-044-506B-15	Sequence 15, Appl
C 734	14	0.9	43280	3	US-08-804-227C-1	Sequence 1, Appl	C 807	13	0.9	48	5	US-09-109-063-18	Sequence 38, Appl
C 735	14	0.9	44377	4	US-08-804-227C-7	Sequence 7, Appl	C 808	13	0.9	57	4	US-08-663-566A-55	Sequence 55, Appl
C 736	14	0.9	44377	4	US-08-804-198-1	Sequence 1, Appl	C 809	13	0.9	57	4	US-08-663-566A-55	Sequence 55, Appl
C 737	14	0.9	49272	2	US-08-614-770A-1	Sequence 1, Appl	C 810	13	0.9	57	4	US-08-288-065A-55	Sequence 55, Appl
C 738	14	0.9	49272	2	US-08-614-770A-1	Sequence 1, Appl	C 810	13	0.9	57	4	US-08-362-240A-55	Sequence 55, Appl
C 739	14	0.9	49377	2	US-08-764-233A-1	Sequence 1, Appl	C 811	13	0.9	57	6	PCT-US95-10245-55	Sequence 55, Appl
C 740	14	0.9	49377	2	US-08-764-233A-1	Sequence 1, Appl	C 812	13	0.9	63	3	US-08-663-566A-55	Sequence 55, Appl
C 741	14	0.9	50341	2	US-08-247-901C-1	Sequence 1, Appl	C 813	13	0.9	63	4	US-08-288-065A-56	Sequence 56, Appl
C 742	14	0.9	50341	4	US-09-075-904-1	Sequence 1, Appl	C 814	13	0.9	63	4	US-08-352-240A-56	Sequence 56, Appl
C 743	14	0.9	246240	3	US-08-724-394A-20	Sequence 20, Appl	C 815	13	0.9	63	6	PCT-US95-10245-56	Sequence 56, Appl
C 744	14	0.9	246240	3	US-08-724-394A-21	Sequence 21, Appl	C 816	13	0.9	67	4	US-08-480-449-17	Sequence 17, Appl
C 745	14	0.9	246240	3	US-08-724-394A-22	Sequence 22, Appl	C 817	13	0.9	67	4	US-08-660-542-17	Sequence 17, Appl
C 746	13	0.9	15	3	US-08-585-684B-2285	Sequence 2285, App	C 818	13	0.9	67	5	US-08-989-107-17	Sequence 17, Appl
C 747	13	0.9	15	3	US-08-585-684B-2285	Sequence 2285, App	C 819	13	0.9	67	5	US-08-388-353-649	Sequence 649, App
C 748	13	0.9	15	3	US-08-585-684B-2285	Sequence 2285, App	C 820	13	0.9	67	5	US-08-488-551B-649	Sequence 649, App
C 749	13	0.9	15	5	US-08-181-664-31	Sequence 31, Appl	C 821	13	0.9	77	1	US-08-442-572-18	Sequence 18, Appl
C 750	13	0.9	18	1	US-08-050-073-209	Sequence 209, App	C 822	13	0.9	77	1	US-08-384-708A-204	Sequence 204, App
C 751	13	0.9	19	1	US-08-050-073-185	Sequence 185, App	C 823	13	0.9	77	2	US-08-361-795-18	Sequence 18, Appl
C 752	13	0.9	20	1	US-08-368-870-5	Sequence 5, Appl	C 824	13	0.9	77	3	US-08-470-939-12	Sequence 12, Appl
C 753	13	0.9	20	1	US-08-471-601-4	Sequence 4, Appl	C 825	13	0.9	77	4	US-08-477-527A-233	Sequence 233, App
C 754	13	0.9	20	1	US-08-474-556-4	Sequence 4, Appl	C 826	13	0.9	77	4	US-08-477-527A-248	Sequence 248, App
C 755	13	0.9	20	1	US-08-351-899-4	Sequence 4, Appl	C 827	13	0.9	77	5	US-08-481-710-233	Sequence 233, App
C 756	13	0.9	20	2	US-08-479-382-4	Sequence 4, Appl	C 828	13	0.9	77	5	US-08-481-710-248	Sequence 248, App
C 756	13	0.9	20	2	US-08-479-382-4	Sequence 4, Appl	C 829	13	0.9	77	6	PCT-US95-05600-101	Sequence 101, App

830	13	0.9	77	6	PCT-US96-09452-2	Sequence 2, Appl1	903	13	0.9	327	3	US-08-429-965-1	Sequence 1, Appl1
831	13	0.9	77	6	PCT-US96-09537-233	Sequence 233, App	904	13	0.9	327	3	US-08-429-965-3	Sequence 3, Appl1
832	13	0.9	77	6	PCT-US96-09537-248	Sequence 248, App	905	13	0.9	327	4	US-08-969-584-1	Sequence 1, Appl1
833	13	0.9	78	4	US-08-477-527A-224	Sequence 224, App	906	13	0.9	327	4	US-08-969-584-3	Sequence 3, Appl1
834	13	0.9	78	4	US-08-481-710-224	Sequence 224, App	907	13	0.9	327	4	US-08-970-179A-1	Sequence 1, Appl1
835	13	0.9	78	6	PCT-US96-09537-224	Sequence 224, App	908	13	0.9	327	4	US-08-970-179A-3	Sequence 3, Appl1
836	13	0.9	81	1	US-08-009-265-9	Sequence 9, Appl1	909	13	0.9	327	4	US-08-969-249A-1	Sequence 1, Appl1
837	13	0.9	83	4	US-08-481-658B-66	Sequence 66, Appl	910	13	0.9	327	5	US-08-969-249A-3	Sequence 3, Appl1
838	13	0.9	83	4	US-08-477-504A-66	Sequence 66, Appl	911	13	0.9	327	5	US-08-969-378-1	Sequence 1, Appl1
839	13	0.9	83	4	US-08-486-756A-66	Sequence 66, Appl	912	13	0.9	327	5	US-08-969-378-3	Sequence 3, Appl1
840	13	0.9	83	4	US-08-485-862B-66	Sequence 66, Appl	913	13	0.9	337	5	US-08-997-080-191	Sequence 5, Appl1
841	13	0.9	83	5	US-08-787-739-66	Sequence 66, Appl	914	13	0.9	337	4	US-08-997-080-191	Sequence 1, App
842	13	0.9	87	1	US-08-433-126A-44	Sequence 44, Appl	915	13	0.9	337	4	US-08-997-362-191	Sequence 191, App
843	13	0.9	87	2	US-08-433-124A-44	Sequence 44, Appl	916	13	0.9	387	1	US-08-458-084-7	Sequence 7, Appl1
844	13	0.9	87	4	US-08-477-527A-85	Sequence 85, Appl	917	13	0.9	387	1	US-08-458-084-8	Sequence 8, Appl1
845	13	0.9	87	4	US-08-477-527A-155	Sequence 155, App	918	13	0.9	387	1	US-08-205-508-7	Sequence 8, Appl1
846	13	0.9	87	5	US-08-481-710-85	Sequence 85, Appl	919	13	0.9	387	2	US-08-171-718-109	Sequence 109, App
847	13	0.9	87	5	US-08-481-710-155	Sequence 155, App	920	13	0.9	387	2	US-08-482-148-4	Sequence 4, Appl1
848	13	0.9	87	6	PCT-US96-06059-44	Sequence 44, Appl	921	13	0.9	387	3	US-08-482-148-5	Sequence 5, Appl1
849	13	0.9	87	6	PCT-US96-09537-85	Sequence 85, Appl	922	13	0.9	387	6	PCT-US95-02944-4	Sequence 4, Appl1
850	13	0.9	87	6	PCT-US96-09537-155	Sequence 155, App	923	13	0.9	387	6	PCT-US95-02944-5	Sequence 5, Appl1
851	13	0.9	94	1	US-08-479-783A-60	Sequence 60, Appl	924	13	0.9	387	6	PCT-US95-02945-7	Sequence 7, Appl1
852	13	0.9	94	1	US-08-479-783A-60	Sequence 60, Appl	925	13	0.9	387	6	PCT-US95-02945-8	Sequence 8, Appl1
853	13	0.9	94	2	US-08-618-693-60	Sequence 60, Appl	926	13	0.9	392	1	US-08-181-271A-101	Sequence 101, App
854	13	0.9	94	6	PCT-US96-08014-149	Sequence 149, App	927	13	0.9	392	1	US-08-449-315-101	Sequence 101, App
855	13	0.9	96	1	US-08-009-265-40	Sequence 40, Appl	928	13	0.9	392	1	US-08-444-803-101	Sequence 101, App
856	13	0.9	108	1	US-07-663-413-19	Sequence 19, Appl	929	13	0.9	392	1	US-08-449-043-101	Sequence 101, App
857	13	0.9	108	1	US-08-055-530-19	Sequence 19, Appl	930	13	0.9	392	2	US-08-456-265A-101	Sequence 101, App
858	13	0.9	120	1	US-08-308-949A-3	Sequence 3, Appl1	931	13	0.9	392	2	US-08-455-416-101	Sequence 101, App
859	13	0.9	122	1	US-08-308-949A-2	Sequence 2, Appl1	932	13	0.9	392	2	US-08-455-244-101	Sequence 101, App
860	13	0.9	125	3	US-08-611-757-81	Sequence 81, Appl	933	13	0.9	392	2	US-08-454-876-101	Sequence 101, App
861	13	0.9	125	3	US-08-611-757-81	Sequence 81, Appl	934	13	0.9	392	2	US-08-457-364-101	Sequence 101, App
862	13	0.9	125	6	PCT-US95-05980-81	Sequence 81, Appl	935	13	0.9	392	3	US-08-456-262-101	Sequence 101, App
863	13	0.9	125	6	PCT-US95-05980-81	Sequence 81, Appl	936	13	0.9	392	3	US-08-456-240-101	Sequence 101, App
864	13	0.9	145	6	PCT-US95-05980-81	Sequence 81, Appl	937	13	0.9	392	3	US-08-455-746-101	Sequence 101, App
865	13	0.9	146	2	US-08-484-570A-54	Sequence 54, Appl	938	13	0.9	392	4	US-08-971-217-101	Sequence 101, App
866	13	0.9	158	4	US-08-687-080-79	Sequence 79, Appl	939	13	0.9	396	1	US-08-458-084-6	Sequence 6, Appl1
867	13	0.9	174	1	US-08-360-841-1	Sequence 1, Appl1	940	13	0.9	396	1	US-08-205-508-6	Sequence 6, Appl1
868	13	0.9	190	6	PCT-US93-06251-14	Sequence 14, Appl	941	13	0.9	396	6	US-08-482-148-3	Sequence 3, Appl1
869	13	0.9	195	1	US-08-044-506B-27	Sequence 27, Appl	942	13	0.9	396	6	PCT-US95-02944-3	Sequence 3, Appl1
870	13	0.9	197	1	US-08-392-678-5	Sequence 5, Appl1	943	13	0.9	412	3	PCT-US95-02945-6	Sequence 6, Appl1
871	13	0.9	197	1	US-07-972-481-3	Sequence 3, Appl1	944	13	0.9	412	3	US-08-807-043-2	Sequence 2, Appl1
872	13	0.9	197	1	US-08-457-304A-5	Sequence 5, Appl1	945	13	0.9	412	4	US-09-127-289-2	Sequence 2, Appl1
873	13	0.9	197	1	US-08-456-701A-5	Sequence 5, Appl1	946	13	0.9	414	4	US-09-067-773-2	Sequence 2, Appl1
874	13	0.9	198	6	PCT-US95-10668-1	Sequence 1, Appl1	947	13	0.9	414	4	US-09-067-773-6	Sequence 6, Appl1
875	13	0.9	198	6	PCT-US95-10668-2	Sequence 2, Appl1	948	13	0.9	417	4	US-08-245-511-11	Sequence 11, Appl
876	13	0.9	214	2	US-08-435-684A-45	Sequence 45, Appl	949	13	0.9	417	4	US-08-600-993A-11	Sequence 11, Appl
877	13	0.9	214	4	US-08-934-877A-45	Sequence 45, Appl	950	13	0.9	423	1	US-08-470-179-149	Sequence 149, App
878	13	0.9	236	7	5179196-4	Patent No. 5179196	951	13	0.9	432	2	US-08-462-216-1	Sequence 2, Appl1
879	13	0.9	236	7	5179196-5	Patent No. 5179196	952	13	0.9	432	2	US-07-915-720D-26	Sequence 26, Appl
880	13	0.9	243	7	5164304-4	Patent No. 5164304	953	13	0.9	432	1	US-08-137-016-13	Sequence 13, Appl1
881	13	0.9	243	7	5164304-5	Patent No. 5164304	954	13	0.9	450	4	US-08-345-321-5	Sequence 5, Appl1
882	13	0.9	243	7	5187075-3	Patent No. 5187075	955	13	0.9	470	3	US-08-967-101-99	Sequence 99, Appl
883	13	0.9	246	7	5482923-1	Patent No. 5482923	956	13	0.9	470	3	US-08-592-541-89	Sequence 89, Appl
884	13	0.9	256	6	PCT-US94-031-75	Sequence 75, Appl	957	13	0.9	474	3	US-08-403-852D-14	Sequence 14, Appl
885	13	0.9	256	6	PCT-US95-16766-28	Sequence 28, Appl	958	13	0.9	477	2	US-08-797-831A-2	Sequence 2, Appl1
886	13	0.9	273	4	US-08-997-080-142	Sequence 142, App	959	13	0.9	477	3	US-08-463-081B-31	Sequence 31, Appl
887	13	0.9	273	4	US-08-997-362-142	Sequence 142, App	960	13	0.9	477	3	US-08-461-379A-31	Sequence 31, Appl
888	13	0.9	277	4	US-08-690-011A-30	Sequence 30, Appl	961	13	0.9	477	3	US-08-462-390B-31	Sequence 31, Appl
889	13	0.9	282	7	5218099-8	Patent No. 5218099	962	13	0.9	477	5	US-08-463-074B-31	Sequence 31, Appl
890	13	0.9	293	4	US-08-752-844-43	Sequence 43, Appl	963	13	0.9	477	5	US-08-465-585C-31	Sequence 31, Appl
891	13	0.9	293	4	US-08-591-196-43	Sequence 43, Appl	964	13	0.9	479	1	US-08-033-857A-1	Sequence 1, Appl1
892	13	0.9	300	3	US-08-446-345-80	Sequence 30, Appl	965	13	0.9	479	1	US-08-374-983A-1	Sequence 1, Appl1
893	13	0.9	302	2	US-08-592-126-84	Sequence 84, Appl	966	13	0.9	480	2	US-07-847-010-2	Sequence 2, Appl1
894	13	0.9	303	4	US-08-865-273-3	Sequence 3, Appl1	967	13	0.9	480	2	US-07-847-010-22	Sequence 22, Appl
895	13	0.9	305	2	US-08-062-472B-5	Sequence 5, Appl1	968	13	0.9	487	2	US-08-453-956-17	Sequence 17, Appl
896	13	0.9	312	1	US-08-248-472A-66	Sequence 66, Appl	969	13	0.9	487	2	US-08-086-631-17	Sequence 17, Appl
897	13	0.9	312	2	US-08-062-472B-4	Sequence 4, Appl1	970	13	0.9	487	6	PCT-US93-08174-17	Sequence 17, Appl
898	13	0.9	317	2	US-08-594-031-73	Sequence 73, Appl1	971	13	0.9	487	6	PCT-US93-08174-17	Sequence 17, Appl
899	13	0.9	318	2	US-08-436-463-9	Sequence 9, Appl1	972	13	0.9	495	1	US-08-133-711-41	Sequence 41, Appl
900	13	0.9	318	2	US-08-024-253-9	Sequence 9, Appl1	973	13	0.9	495	2	US-08-406-248-1	Sequence 1, Appl1
901	13	0.9	327	2	US-08-638-610-1	Sequence 1, Appl1	974	13	0.9	501	2	US-08-470-892-1	Sequence 1, Appl1
902	13	0.9	327	2	US-08-638-610-3	Sequence 3, Appl1	975	13	0.9	504	1	US-07-661-610C-15	Sequence 15, Appl

c 976 13 0.9 508 4 US-08-975-316-26 Sequence 26, Appl
977 13 0.9 515 1 US-08-543-238-12 Sequence 12, Appl
978 13 0.9 515 1 US-08-420-526-12 Sequence 12, Appl
979 13 0.9 530 2 US-08-581-529B-5 Sequence 5, Appl
980 13 0.9 530 6 PCT-US94-07762-5 Sequence 5, Appl
981 13 0.9 531 6 PCT-US95-04583-1 Sequence 1, Appl
982 13 0.9 553 1 US-08-044-506B-22 Sequence 22, Appl
983 13 0.9 573 3 US-08-290-665A-136 Sequence 136, Appl
984 13 0.9 573 6 PCT-US95-10398-136 Sequence 136, Appl
985 13 0.9 576 2 US-08-453-956-16 Sequence 16, Appl
986 13 0.9 576 2 US-08-086-631-16 Sequence 16, Appl
987 13 0.9 576 4 US-08-452-930-16 Sequence 16, Appl
988 13 0.9 576 6 PCT-US93-08174-16 Sequence 16, Appl
989 13 0.9 577 5 US-08-938-548B-21 Sequence 21, Appl
990 13 0.9 585 5 US-08-938-548B-5 Sequence 5, Appl
991 13 0.9 591 4 US-09-156-979-1 Sequence 1, Appl
992 13 0.9 607 4 US-08-975-316-23 Sequence 1, Appl
993 13 0.9 612 3 US-08-308-952-9 Sequence 23, Appl
994 13 0.9 617 4 US-08-481-658B-47 Sequence 9, Appl
995 13 0.9 617 4 US-08-477-504A-47 Sequence 47, Appl
996 13 0.9 617 4 US-08-486-756A-47 Sequence 47, Appl
997 13 0.9 617 4 US-08-485-862B-47 Sequence 47, Appl
998 13 0.9 617 5 US-08-787-739-47 Sequence 47, Appl
999 13 0.9 618 2 US-08-800-751-3 Sequence 3, Appl
c1000 13 0.9 618 4 US-08-990-818-3 Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-554-659-7/c
Sequence 7, Application US/08554659
Patent No. 5767261
GENERAL INFORMATION:
APPLICANT: Wingate, Vincent
TITLE OF INVENTION: LEPIDOPTERAN GABA GATED CHLORIDE CHANNEL
TITLE OF INVENTION: AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/554,659
FILING DATE:
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9732
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ. ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 128 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-554-659-7

Query Match 1.3%; Score 19; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 TCGAGCGCAAGACCATGAC 323
Db 127 TCGAGCGCAAGACCATGAC 109

RESULT 2
US-08-140-104A-4/c
Sequence 4, Application US/08140104A
Patent No. 5585255
GENERAL INFORMATION:
APPLICANT: Tsukada, Yogi
APPLICANT: Tazuke, Yasuhiko
APPLICANT: Okada, Shigenori
APPLICANT: Adachi, Kenichi
TITLE OF INVENTION: BILE ACID SULFATASE GENE, PLASMID
TITLE OF INVENTION: CONTAINING SAID GENE AND METHOD OF PRODUCING BILE ACID
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/140,104A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP PCT/JP93/00244
FILING DATE: 26-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E.
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: SAEGU3.001AUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
INFORMATION FOR SEQ. ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1488 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: BILE ACID SULFATASE
FEATURE:
NAME/KEY: CDS
LOCATION: 106..1509
US-08-140-104A-4

Query Match 1.1%; Score 17; DB 1; Length 1488;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1310 CCAAGCGACCGCTTC 1326
Db 206 CCAAGCGACCGCTTC 190

RESULT 3
US-08-140-104A-3/C
; Sequence 3, Application US/08140104A
; Patent No. 5585255
; GENERAL INFORMATION:
; APPLICANT: Tsukada, Yogi
; APPLICANT: Tazuke, Yasuhiko
; APPLICANT: Okada, Shigenori
; APPLICANT: Adachi, Kenichi
; TITLE OF INVENTION: BILE ACID SULFATASE GENE, PLASMID
; TITLE OF INVENTION: CONTAINING SAID GENE AND METHOD OF PRODUCING BILE ACID
; TITLE OF INVENTION: SULFATE SULFATASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/140,104A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP PCT/JP93/00244
; FILING DATE: 26-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E.
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: SAE03.001AUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1497 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: BILE ACID SULFATASE
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 106..1509
US-08-140-104A-3

Query Match 1.1%; Score 17; DB 1; Length 1497;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4
US-08-850-910A-17
; Sequence 17, Application US/08850910A
; Patent No. 5948761
; GENERAL INFORMATION:
; APPLICANT: SELHAMER, J.J.
; APPLICANT: SELHAMER, J.J.
; APPLICANT: SCARBOROUGH, R.M.
; TITLE OF INVENTION: RECOMBINANT TECHNIQUES FOR
; TITLE OF INVENTION: PRODUCTION OF BRAIN NATRIURETIC PEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER, LLP
; STREET: 2000 Pennsylvania Avenue, NW, Suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850,910A
; FILING DATE: 05-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/477,226
; FILING DATE: 08-FEB-1990
; APPLICATION NUMBER: 07/299,880
; FILING DATE: 19-JAN-1989
; APPLICATION NUMBER: 07/206,470
; FILING DATE: 14-JUN-1988
; APPLICATION NUMBER: 07/200,383
; FILING DATE: 31-MAY-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 219002025212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; TELEFAX: 202-822-0168
; TELEX:
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1504 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 100...630
; OTHER INFORMATION:
US-08-850-910A-17

Query Match 1.1%; Score 17; DB 4; Length 1504;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 39 GTTGCTGGAGGCGC 55
DB 251 GTTGCTGGAGGCGC 267
US-08-850-910A-38
; Sequence 38, Application US/08850910A
; Patent No. 5948761
; GENERAL INFORMATION:
; APPLICANT: SELHAMER, J.J.
; APPLICANT: SELHAMER, J.J.
; APPLICANT: SCARBOROUGH, R.M.
; TITLE OF INVENTION: RECOMBINANT TECHNIQUES FOR
; TITLE OF INVENTION: PRODUCTION OF BRAIN NATRIURETIC PEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER, LLP

STREET: 2000 Pennsylvania Avenue, NW, Suite 5500
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: Fastseq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/850, 910A
FILING DATE: 05-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/477, 226
FILING DATE: 08-FEB-1990
APPLICATION NUMBER: 07/299, 880
FILING DATE: 19-JAN-1989
APPLICATION NUMBER: 07/206, 470
FILING DATE: 14-JUN-1988
APPLICATION NUMBER: 07/200, 383
FILING DATE: 31-MAY-1988
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H
REGISTRATION NUMBER: 29, 959
REFERENCE/DOCKET NUMBER: 219002025212
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-887-1500
TELEFAX: 202-822-0168
TELEX:
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1507 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 100...222
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: 469...723
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: 1281...1292
OTHER INFORMATION:
US-08-850-910A-38

```

Query Match          1.1%; Score 17; DB 4; Length 1507;
Best Local Similarity 100.0%; Pred No. 24;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      39 GTTGCGTGGAGGCGC 55
        |||
Db       251 GTTGCGTGGAGGCGC 267

RESULT    6
US-08-140-104A-1/c
; Sequence 1, Application US/08140104A
; Patent No. 5585255
; GENERAL INFORMATION:
; APPLICANT: Tsukada, Yogi
; APPLICANT: Tazuke, Yasuhiko
; APPLICANT: Okada, Shigenori
; APPLICANT: Adachi, Kenichi
; TITLE OF INVENTION: BILE ACID SULFATASE GENE, PLASMID
; TITLE OF INVENTION: CONTAINING SAID GENE AND METHOD OF PRODUCING BILE ACID
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: Knobbe, Martens Olson and Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/140,104A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP PCT/JP93/00244
FILING DATE: 26-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E.
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: SAC03.001AUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
INFORMATION FOR SEQ. ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1509 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: BILE ACID SULFATASE
FEATURE:
NAME/KEY: CDS
LOCATION: 106..1509
US-08-140-104A-1

	Query Match	1.1%	Score 17	DB 1	Length 1509
	Best Local Similarity	100.0%	Pred. No. 24		
Matches	17	Conservative	0	Mismatches	0
				Indels	0
				Gaps	0
QY	1310	CCAAAGCCACCGCCTTC	1326		
Db	227	CCAAAGCCACCGCCTTC	211		

RESULT 7
 5177197-31
 Patent NO. 5177197
 APPLICANT: KANZAKI, TETSUO; OLOFSSON, ANDERS; MOREN, ANITA;
 WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSON-WELSH;
 LENA; HELDIN, CARL-HENRIC
 TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
 HUMAN TRANSFORMING GROWTH FACTOR-BETA1-BINDING PROTEIN
 NUMBER OF SEQUENCES: 53
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/487,343
 FILING DATE: 27-FEB-1990
 SEQ ID NO.: 31:
 LENGTH: 5089
 5177197-31

Query Match	1.1%	Score 17	DB 7	Length 5089
Best Local Similarity	100.0%	Pred. No. 23		
Matches 17	Conservative 0	Mismatches 0	Indels 0	Gaps 0
0y	249	TCAGGTGCACCTGTGA	265	

Db 2411 tcagtgccactgtgga 2427

|||||
RESULT 8
US-08-030-731A-40/c
; Sequence 40, Application US/08030731A
; Patent No. 5426036
; GENERAL INFORMATION:
; APPLICANT: Koller, Klaus-Peter
; APPLICANT: Riess, Guenther Johannes
; APPLICANT: Uhlmann, Eugen
; APPLICANT: Wellmeier, Holger
; TITLE OF INVENTION: Processes for the Preparation of Foreign
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/030,731A
; FILING DATE: 12-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/189,840
; FILING DATE: 03-MAY-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/430,622
; FILING DATE: 01-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/687,610
; FILING DATE: 19-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/735,757
; FILING DATE: 29-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 37 14 866.4
; FILING DATE: 05-MAY-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 38 37 273.8
; FILING DATE: 03-NOV-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 39 27 449.7
; FILING DATE: 19-AUG-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 40 12 818.0
; FILING DATE: 21-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Kirschner Michael K.
; REGISTRATION NUMBER: 34,851
; REFERENCE/DOCKET NUMBER: 02481-0593-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4000
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-030-731A-40

Query Match 1.1%; Score 16; DB 1; Length 225;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 349 GCCGATGACCGCTCG 364
|||||
Db 186 GCCGATGACCGCTCG 171

RESULT 9
US-08-716-942-23
; Sequence 23, Application US/08716942
; Patent No. 5849491
; GENERAL INFORMATION:
; APPLICANT: Terragen Diversity Inc.
; APPLICANT: Radomski, Christopher C. A.
; APPLICANT: Seow, Kah Tong
; APPLICANT: Warren, R. Anthony J.
; TITLE OF INVENTION: METHOD FOR ISOLATING XYLANASE GENE
; TITLE OF INVENTION: SEQUENCES FROM SOIL DNA, COMPOSITIONS USEFUL IN SUCH METHOD
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oppedahl & Larson
; STREET: 1992 Commerce Street, Suite 309
; CITY: Yorktown Heights
; STATE: NY
; COUNTRY: USA
; ZIP: 10598-4412
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS 5.0
; SOFTWARE: Word Perfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/716,942
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/004,157
; FILING DATE: 20-Sep-95
; ATTORNEY/AGENT INFORMATION:
; NAME: Marina T. Larson
; REGISTRATION NUMBER: 32,038
; REFERENCE/DOCKET NUMBER: TERR-P-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 245-3252
; TELEFAX: (914) 962-4330
; TELEX:
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 285
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Cellulomonas fimi
; FEATURE:
; NAME/KEY: sequence of internal fragments of xylanase gene
US-08-716-942-23

Query Match 1.1%; Score 16; DB 3; Length 285;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 407 TCGAGGCGAAGTTCGC 422

Db 113 TCGAGGCCAGGTCCG 128

RESULT 10
PCT-US96-05320A-1376/c

Sequence 1376, Application PC/TUS9605320A

GENERAL INFORMATION:

APPLICANT: Human Genome Sciences

APPLICANT: 9410 Key West Avenue

APPLICANT: Rockville, MD 20850

APPLICANT: United States of America

APPLICANT: Johns Hopkins University

APPLICANT: 720 Rutland Avenue

APPLICANT: Baltimore, MD 21205

APPLICANT: United States of America

APPLICANT: Mark D. Adams

APPLICANT: Owen White

APPLICANT: Hamilton O. Smith

APPLICANT: J. Craig Venter

TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome,

NUMBER OF SEQUENCES: 48

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20003-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 Inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/05320A

FILING DATE: Apr1122, 1996

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/476,102

FILING DATE: June 7, 1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/487,429

FILING DATE: June 7, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Eric K. Steffe

REGISTRATION NUMBER: 36,688

REFERENCE/DOCKET NUMBER: 1488, 014PC01

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 1376:

SEQUENCE CHARACTERISTICS:

LENGTH: 729 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

PCT-US96-05320A-1376

Query Match 1.1%; Score 16; DB 6; Length 729;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 235 TCGTTCCTTCGGCGC 220

RESULT 11

5512669-1

Patent No. 5512669

APPLICANT: PEOPLES, OLIVER P.; SINSKEY, ANTHONY J.

TITLE OF INVENTION: GENE ENCODING BACTERIAL ACETOACETYL-COA

REDUCTASE

NUMBER OF SEQUENCES: 7

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/297,667

FILING DATE: 29-AUG-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 124,570

FILING DATE: 20-SEP-1993

APPLICATION NUMBER: 944,488

FILING DATE: 03-NOV-1992

APPLICATION NUMBER: 566,535

FILING DATE: 13-AUG-1990

APPLICATION NUMBER: 67,695

FILING DATE: 29-JUN-1987

SEQ ID NO:1:

LENGTH: 1524

5512669-1

Query Match 1.1%; Score 16; DB 7; Length 1524;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 815 CGCCTCGCGTCCGCGC 830

RESULT 12

5229279-1

Patent No. 5229279

APPLICANT: PEOPLES, OLIVER P.; SINSKEY, ANTHONY J.

TITLE OF INVENTION: METHOD FOR PRODUCING NOVEL POLYESTER

BIOPOLYMERS

NUMBER OF SEQUENCES: 8

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/556,535

FILING DATE: 13-AUG-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 67,695

FILING DATE: 29-AUG-1987

SEQ ID NO:1:

LENGTH: 1525

5229279-1

Query Match 1.1%; Score 16; DB 7; Length 1525;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 816 CGCCTCGCGTCCGCGC 831

RESULT 13

5352575-8

Patent No. 5352575

APPLICANT: PETROVSKIS, ERIK A.; POST, LEONARD E.; TIMMINS, JAMES G.

TITLE OF INVENTION: PSEUDORABIES VIRUS PROTEIN

NUMBER OF SEQUENCES: 12

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/513,282

FILING DATE: 20-APR-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 100,817

FILING DATE: 29-JUN-1987

APPLICATION NUMBER: 886,260

FILING DATE: 16-JUL-1986

APPLICATION NUMBER: 784,787

FILING DATE: 04-OCT-1985

APPLICATION NUMBER: 801,799

FILING DATE: 26-NOV-1985
APPLICATION NUMBER: 844,113
FILING DATE: 26-MAR-1986
SEQ ID NO:8
LENGTH: 1734
5352575-8

Query Match 1.1%; Score 16; DB 7; Length 1734;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1065 CGAGGACTACTACGAC 1080
|||||
DB 1452 CGAGGACTACTACGAC 1467

RESULT 14

US-08-442-134A-1
Sequence 1, Application US/08442134A
Patent No. 5596088

GENERAL INFORMATION:
APPLICANT: Boucher, Richard C.
APPLICANT: Weisman, Gary A.
APPLICANT: Turner, John T.
APPLICANT: Harden, Thomas K.
APPLICANT: Parr, Claude E.
APPLICANT: Sullivan, Daniel M.
APPLICANT: Erb, Laura
APPLICANT: Lustig, Kevin D.
TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and
TITLE OF INVENTION: Null Cells Expressing P2U Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell, Seltzer, Park & Gibson
STREET: Post Office Drawer 34009
CITY: Charlotte
STATE: No. 5596088th Carolina
COUNTRY: USA
ZIP: 28234

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/442,134A
FILING DATE: 16-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-71A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1842 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 57..1181
US-08-442-134A-1

Query Match 1.1%; Score 16; DB 1; Length 1842;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 850 GTGCGCTTCCTCTCT 865
|||||
DB 381 GTGCGCTTCCTCTCT 396

RESULT 15

US-08-444-581B-1
Sequence 1, Application US/08444581B
Patent No. 5607836

GENERAL INFORMATION:
APPLICANT: Boucher, Richard C.
APPLICANT: Weisman, Gary A.
APPLICANT: Turner, John T.
APPLICANT: Harden, Thomas K.
APPLICANT: Parr, Claude E.
APPLICANT: Sullivan, Daniel M.
APPLICANT: Erb, Laura
APPLICANT: Lustig, Kevin D.
TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and
TITLE OF INVENTION: Null Cells Expressing P2U Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell, Seltzer, Park & Gibson
STREET: Post Office Drawer 34009
CITY: Charlotte
STATE: No. 5607836th Carolina
COUNTRY: USA
ZIP: 28234

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,581B
FILING DATE: 19-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/442,134
FILING DATE: 16-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-71A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1842 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 57..1181
US-08-444-581B-1

Query Match 1.1%; Score 16; DB 1; Length 1842;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 850 GTGCGCTTCCTCTCT 865
|||||
DB 381 GTGCGCTTCCTCTCT 396

Search completed: May 11, 2000, 00:35:21
Job time: 7735 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2000, 22:04:09 ; Search time 495.57 Seconds
(without alignments)
12194.757 Million cell updates/sec

Title: US-09-080-127-1
Perfect score: 1491
Sequence: 1 ATGAGGTCCGCTTTGTGGGC.....ATTCTCAAGTCGAGCATAG 1491

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4857316 seqs, 202611650 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est34:*
60: gb_est35:*
61: gb_est36:*
62: gb_est37:*
63: gb_est38:*
64: gb_est39:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	1.3	286	40	AA985688	AA985688 uc13a03.y
2	20	1.3	491	107	AQ444631	AQ444631 GSSC0139
3	20	1.3	713	102	AQ162320	AQ162320 mgxb00121
4	19	1.3	392	92	AA556143	AA556143 GSSC0594
5	19	1.3	392	92	AA556143	AA556143 TENF0072
6	19	1.3	401	92	AQ094862	AQ094862 GSSC0802
7	19	1.3	401	92	AQ094862	AQ094862 GSSC0802
8	19	1.3	428	79	AA283695	AA283695 LCL1_222.D
9	19	1.3	529	109	AA068243	AA068243 HS_3435.A
10	19	1.3	367	84	B62537	B62537 T22C147E.TA
11	19	1.3	572	50	AI667964	AI667964 TENG0802
12	19	1.3	587	50	AI667964	AI667964 TENG0802
13	18	1.2	235	47	AI527261	AI527261 u150f05.x
14	18	1.2	300	35	C32880	C32880 C32880.YuJ1
15	18	1.2	300	62	AV175791	AV175791 AV175791
16	18	1.2	300	62	AV199192	AV199192 AV199192
17	18	1.2	349	84	B87717	B87717 RPTC111-30D1
18	18	1.2	354	29	AA182748	AA182748 zP40f10.s
19	18	1.2	357	103	AA021847	AA021847 HS_2001.B
20	18	1.2	361	43	AI213825	AI213825 z5905a1.f
21	18	1.2	369	79	AA294468	AA294468 UT-H-B12
22	18	1.2	383	47	AI536432	AI536432 vU28f12.x
23	18	1.2	402	43	C97509	C97509 C97509.Rice
24	18	1.2	405	43	AI213824	AI213824 z5905a1.f
25	18	1.2	433	90	AA080647	AA080647 HS_3159.A
26	18	1.2	441	105	AA031349	AA031349 RPTC111-10
27	18	1.2	451	107	AA098464	AA098464 HS_5088.B
28	18	1.2	459	85	AA0720719	AA0720719 HS_5546.B
29	18	1.2	485	74	AA200193	AA200193 d12g08.Y
30	18	1.2	491	36	AA644917	AA644917 v84d007.r
31	18	1.2	536	25	N63397	N63397 yf60a05.s1
32	18	1.2	541	49	AI653733	AI653733 wB36g03.x
33	18	1.2	549	37	AA681631	AA681631 v43b12.s
34	18	1.2	549	46	AI455750	AI455750 ID35425.s
35	18	1.2	556	102	AA0154304	AA0154304 HS_3006.A
36	18	1.2	559	105	AA0261972	AA0261972 CITR1-El-
37	18	1.2	571	35	AA546920	AA546920 v467f12.s
38	18	1.2	581	36	AA623971	AA623971 v474g10.s
39	18	1.2	584	36	AA623914	AA623914 v474g10.s
40	18	1.2	597	39	AA891236	AA891236 EST195039
41	18	1.2	612	84	B89598	B89598 CIT-HSP-216
42	18	1.2	628	106	AA0388027	AA0388027 RPTC111-15
43	18	1.2	638	46	AI405569	AI405569 G125609.s
44	18	1.2	681	102	AA010912	AA010912 CIT-HSP-2
45	18	1.2	709	105	AA028192	AA028192 RPTC111-79
46	18	1.2	722	108	AA0578172	AA0578172 nbx0083N
47	18	1.2	738	91	AA0892553	AA0892553 HS_3174.B
48	17	1.1	134	64	AA061871	AA061871 687008B10
49	17	1.1	157	102	AA0142506	AA0142506 HS_3090.A
50	17	1.1	205	59	AA079937	AA079937 AA079937
51	17	1.1	228	28	AA090323	AA090323 y0045.seq
52	17	1.1	239	61	AAV158026	AAV158026 AAV158026
53	17	1.1	242	43	AAI195359	AAI195359 u119f05.x
54	17	1.1	257	31	AA324374	AA324374 EST27186
55	17	1.1	272	36	AA659860	AA659860 nu85a11.s
56	17	1.1	289	20	D31182	D31182 HUML12818.H
57	17	1.1	300	27	C10818	C10818 C10818.YuJ1
58	17	1.1	300	35	C31906	C31906 C31906.YuJ1
59	17	1.1	300	35	C34691	C34691 C34691.YuJ1
60	17	1.1	300	35	C53385	C53385 C53385.YuJ1
61	17	1.1	300	62	C54852	C54852 C54852.YuJ1
62	17	1.1	300	62	AAV200170	AAV200170 AV200170
63	17	1.1	300	62	AAV200408	AAV200408 AV200408
64	17	1.1	312	27	W82411	W82411 me80h04.r1
65	17	1.1	325	40	C93601	C93601 C93601.Rice
66	17	1.1	332	45	AA1356176	AA1356176 gY65d05.x
67	17	1.1	337	24	D66138	D66138 CELK092EXR
68	17	1.1	339	27	C05284	C05284 C05284.Huma
69	17	1.1	344	21	RA04004	RA04004 PK13g12.s1
70	17	1.1	345	20	Z37191	Z37191 ATTS3934.St
71	17	1.1	348	21	D46363	D46363 R1CS10982A
72	17	1.1	352	51	AI750884	AI750884 cm06c04.Y
73	17	1.1	355	30	AA230746	AA230746 mw66a10.r
74	17	1.1	359	49	AI631224	AI631224 t558c05.x
75	17	1.1	360	20	D36406	D36406 CELK0326F
76	17	1.1	364	91	AA0852373	AA0852373 LMAJFV1_1
77	17	1.1	367	91	AA0846241	AA0846241 LMAJFV1_1
78	17	1.1	383	61	AI842270	AI842270 UT-M-A11-
79	17	1.1	385	94	AA025539	AA025539 BP(H)1613
80	17	1.1	397	34	AA467752	AA467752 nc74g03.s
81	17	1.1	402	20	D33792	D33792 CELK036CX
82	17	1.1	403	26	W58216	W58216 zd19e03.s1
83	17	1.1	403	33	AA410892	AA410892 zV03c12.s
84	17	1.1	408	44	AI318865	AI318865 a3b10m.f
85	17	1.1	408	62	AI1902170	AI1902170 IL-EP002-
86	17	1.1	412	79	AI135525	AI135525 DKFZP762L
87	17	1.1	413	28	C20050	C20050 C20050.Rice
88	17	1.1	415	20	D32324	D32324 CELK010GXR
89	17	1.1	418	23	H22366	H22366 yN68g06.s1
90	17	1.1	421	28	C19569	C19569 C19569.Rice
91	17	1.1	422	80	AA379207	AA379207 MR0-HT024
92	17	1.1	423	103	AA0175722	AA0175722 HS_3211.A
93	17	1.1	425	34	AA523136	AA523136 m162f02.s
94	17	1.1	426	20	T01778	T01778 WEST02499.E
95	17	1.1	428	42	AI158277	AI158277 ud23h01.r
96	17	1.1	428	45	AI362891	AI362891 gY81a04.x
97	17	1.1	428	82	FR0008868	FR0008868 F.rubrip
98	17	1.1	430	34	AA029053	AA029053 RPTC111-39
99	17	1.1	431	37	AA692913	AA692913 v56c09.s
100	17	1.1	436	94	AA0029529	AA0029529 RPTC111-39
101	17	1.1	439	47	AI485476	AI485476 EST243797
102	17	1.1	445	108	AA054667	AA054667 CITR1-El-
103	17	1.1	447	27	AA026724	AA026724 z693c01.r
104	17	1.1	450	108	AA0546614	AA0546614 CITR1-El-
105	17	1.1	451	26	W56608	W56608 zd15g11.r1
106	17	1.1	453	37	AA659095	AA659095 GM02130.s
107	17	1.1	454	26	W19193	W19193 zB98c01.s1
108	17	1.1	454	103	AA108609	AA108609 HS_3171.B
109	17	1.1	464	28	AA102547	AA102547 zN41f03.s
110	17	1.1	464	106	AA0354179	AA0354179 CITR1-El-
111	17	1.1	468	40	AA949334	AA949334 LD28220.s
112	17	1.1	468	92	AA0913453	AA0913453 nbcd0041L
113	17	1.1	478	84	B17464	B17464 345L11.TPB
114	17	1.1	479	41	AI390209	AI390209 mw66a10.Y
115	17	1.1	479	85	AA0695379	AA0695379 HS_2148.A
116	17	1.1	486	60	AI804338	AI804338 C556D05.x
117	17	1.1	489	103	AA0177682	AA0177682 HS_2209.B
118	17	1.1	492	102	AA0124228	AA0124228 HS_3122.B
119	17	1.1	493	45	AI388176	AI388176 GH18986.s
120	17	1.1	495	60	AI789193	AI789193 uk52p05.Y
121	17	1.1	495	108	AA0530614	AA0530614 RPTC111-3
122	17	1.1	498	48	AI597335	AI597335 m189f09.Y
123	17	1.1	511	107	AA0510092	AA0510092 nbx0094F
124	17	1.1	515	103	AA0217086	AA0217086 HS_3252.B
125	17	1.1	516	35	U92741	U92741 EST y Boorh
126	17	1.1	516	107	AA0433065	AA0433065 HS_5063.A
127	17	1.1	526	25	N96166	N96166 21622.Lambda
128	17	1.1	535	105	AA0322277	AA0322277 RPTC111-10
129	17	1.1	536	82	FR0008889	FR0008889 F.rubrip
130	17	1.1	537	30	AA272653	AA272653 v476g02.r
131	17	1.1	540	49	AI658229	AI658229 C200403.Y
132	17	1.1	541	109	AA0829183	AA0829183 HS_5168.B
133	17	1.1	546	91	AA0829183	AA0829183 HS_5112.B
134	17	1.1	547	60	AI783400	AI783400 614011D06
135	17	1.1	548	34	AA530645	AA530645 vJ44a06.r
136	17	1.1	552	74	AAW216240	AAW216240 687043F09
137	17	1.1	553	45	AI1179798	AI1179798 EST223358
138	17	1.1	556	34	AA530564	AA530564 vJ48g09.r
139	17	1.1	557	85	AA0735484	AA0735484 HS_3033.B
140	17	1.1	559	109	AA0634406	AA0634406 RPTC111-4
141	17	1.1	563	109	AA0603555	AA0603555 HS_2126.A
142	17	1.1	568	43	AI238714	AI238714 GH14813.s
143	17	1.1	573	63	AA011655	AA011655 614011D06

C 144	17	1.1	573	82	FR0031389
C 145	17	1.1	579	36	AA639254
C 146	17	1.1	580	61	A1828997 wj48902.x
C 147	17	1.1	587	62	A1898461 ESTr267904
C 148	17	1.1	588	46	A1425752 me80h04.y
C 149	17	1.1	596	35	AA540997 LD20736.5
C 150	17	1.1	597	91	AO893611 HS_3142.A
C 151	17	1.1	603	84	B29790 t17C16TfB.T
C 152	17	1.1	616	107	AQ492299 V119c11.m
C 153	17	1.1	619	23	H19692 yns7c05.r.1
C 154	17	1.1	619	109	AO628987 T120004G
C 155	17	1.1	628	105	AO289397 nbxD0033C
C 156	17	1.1	630	40	AA967253 v270c11.t
C 157	17	1.1	635	80	AM376987 IL3-CHro21
C 158	17	1.1	642	46	A1402364 GH21562.5
C 159	17	1.1	646	50	A1667896 TENG0875
C 160	17	1.1	647	42	A1113698 GH10129.5
C 161	17	1.1	671	61	AL043786 DKFZpa34P
C 162	17	1.1	674	43	AU032860 AU032860
C 163	17	1.1	674	61	A1819966 wJ8gq11.x
C 164	17	1.1	679	92	AQ057154 LERAO16Tr
C 165	17	1.1	691	28	AA117324 mm26b11.t
C 166	17	1.1	691	106	AQ385880 RPrC11-15
C 167	17	1.1	713	108	AQ574094 nbxD0085M
C 168	17	1.1	732	90	AO751615 HS_-5576_B
C 169	17	1.1	734	74	AM195468 xn38902.x
C 170	17	1.1	743	44	A1325033 mc042d10.x
C 171	17	1.1	747	92	AQ012212 nbep0034B
C 172	17	1.1	753	41	A1055219 coau0003G
C 173	17	1.1	768	79	AW:67812 ESTR305940
C 174	17	1.1	768	105	AQ290643 nbxb0037M
C 175	17	1.1	779	50	AU067540 AU067540
C 176	17	1.1	788	90	AO788133 HS_-3135.A
C 177	17	1.1	849	108	AQ579204 nbxb0084A
C 178	17	1.1	897	40	AA979887 MEST2-H2.
C 179	17	1.1	916	92	AO914868 npeb0050K
C 180	17	1.1	920	91	AO867991 npeb0018F
C 181	17	1.1	1029	91	AO896897 HS_-5440_A
C 182	17	1.1	1201	83	CNSO16EP
C 183	16	1.1	96	30	AA230257 nc36e08.s
C 184	16	1.1	133	45	A1351099 qfz3bD1.x
C 185	16	1.1	140	34	AA476862 zu29c02.t
C 186	16	1.1	142	51	A1720816 as67a12.x
C 187	16	1.1	170	22	R58064 F9083 Fetat
C 188	16	1.1	174	79	AAW99992 xs55c07.x
C 189	16	1.1	174	50	A1689842 kx31d01.x
C 190	16	1.1	180	36	C61879
C 191	16	1.1	182	74	AA177707 RC1-CtO19
C 192	16	1.1	188	28	AA075203 zmb3e10.s
C 193	16	1.1	188	37	AA728921 nv01d08.s
C 194	16	1.1	193	59	AV086472 AV086472
C 195	16	1.1	197	64	AM099678 sd29f08.y
C 196	16	1.1	204	39	AA866235 vq49c12.t
C 197	16	1.1	206	28	AA116399 mq08a10.t
C 198	16	1.1	210	106	AQ428982 CTBTI-EI-
C 199	16	1.1	210	44	A1258632 LP01866.5
C 200	16	1.1	210	69	AA121785 UI-M-BH2.
C 201	16	1.1	214	39	C23962
C 202	16	1.1	215	59	A1774355
C 203	16	1.1	218	28	AA075202 zmb3e10.t
C 204	16	1.1	219	59	AV112281
C 205	16	1.1	223	70	AV238601
C 206	16	1.1	224	32	AA364754
C 207	16	1.1	225	80	AA325940
C 208	16	1.1	228	38	AA749508
C 209	16	1.1	229	69	AA124463 UI-M-BH2.
C 210	16	1.1	233	62	AA1891305
C 211	16	1.1	234	72	AAV291585
C 212	16	1.1	236	72	AAV287275
C 213	16	1.1	237	71	AAV284176
C 214	16	1.1	238	70	AAV231901
C 215	16	1.1	241	72	AAV308443
C 216	16	1.1	242	59	AAV73366
C 217	16	1.1	217	60	AAV154811
C 218	16	1.1	218	70	AAV232528
C 219	16	1.1	246	80	AAV313954
C 220	16	1.1	252	32	AA371266
C 221	16	1.1	253	80	AAV315809
C 222	16	1.1	256	36	AA649282
C 223	16	1.1	260	62	AA1931651
C 224	16	1.1	260	94	AAQ070838
C 225	16	1.1	262	71	AAV269018
C 226	16	1.1	264	20	AA26472
C 227	16	1.1	266	31	AA299640
C 228	16	1.1	271	28	AA060437
C 229	16	1.1	274	73	AAV350986
C 230	16	1.1	278	60	AAV121986
C 231	16	1.1	278	92	AAQ365020
C 232	16	1.1	282	39	AA881991
C 233	16	1.1	283	28	AA117107
C 234	16	1.1	285	80	AA346872
C 235	16	1.1	288	70	AAV221029
C 236	16	1.1	290	74	AAW227340
C 237	16	1.1	293	106	AAQ351844
C 238	16	1.1	295	46	AA143707
C 239	16	1.1	296	23	AA32340
C 240	16	1.1	296	72	AAV10206
C 241	16	1.1	298	23	AA32213
C 242	16	1.1	299	59	AAV075156
C 243	16	1.1	301	41	AAI008539
C 244	16	1.1	303	81	AAV405805
C 245	16	1.1	309	21	AAV6777
C 246	16	1.1	312	38	AAV59339
C 247	16	1.1	313	24	AAV57542
C 248	16	1.1	313	80	AAV15758
C 249	16	1.1	314	39	AA840035
C 250	16	1.1	316	61	AA1872862
C 251	16	1.1	317	40	AAV000127
C 252	16	1.1	318	30	AA198514
C 253	16	1.1	318	38	AA5440185
C 254	16	1.1	320	21	AAV05984
C 255	16	1.1	320	24	AAV67764
C 256	16	1.1	321	67	AAV582735
C 257	16	1.1	324	20	AAV79026
C 258	16	1.1	324	24	AAV81802
C 259	16	1.1	325	64	AAV054228
C 260	16	1.1	330	36	AAV643268
C 261	16	1.1	331	20	AAV19215
C 262	16	1.1	331	45	AAV1396249
C 263	16	1.1	332	41	AAV1037375
C 264	16	1.1	334	44	AAV054287
C 265	16	1.1	335	63	AAV1232668
C 266	16	1.1	336	48	AAV1599212
C 267	16	1.1	338	35	AAV566959
C 268	16	1.1	342	38	AAV746892
C 269	16	1.1	342	44	AAV1305047
C 270	16	1.1	346	25	AAV19944
C 271	16	1.1	346	46	AAV956123
C 272	16	1.1	346	48	AAV1626633
C 273	16	1.1	347	82	AAV00027402
C 274	16	1.1	351	37	AAV76666
C 275	16	1.1	351	81	AAV431394
C 276	16	1.1	355	35	AAV548891
C 277	16	1.1	355	45	AAV005351
C 278	16	1.1	358	79	AAV252271
C 279	16	1.1	359	43	AAV030546
C 280	16	1.1	360	24	AAV73757
C 281	16	1.1	360	24	AAV73559
C 282	16	1.1	360	35	AAV41499
C 283	16	1.1	360	35	AAV48418
C 284	16	1.1	360	62	AAV188970
C 285	16	1.1	360	80	AAV366381
C 286	16	1.1	361	28	AAV11326
C 287	16	1.1	361	49	AAV1647703
C 288	16	1.1	362	84	AAV63632
C 289	16	1.1	365	20	AAV15553
C 290	16	1.1	365	20	AAV15553
C 291	16	1.1	365	20	AAV15553
C 292	16	1.1	365	20	AAV15553
C 293	16	1.1	365	20	AAV15553
C 294	16	1.1	365	20	AAV15553
C 295	16	1.1	365	20	AAV15553
C 296	16	1.1	365	20	AAV15553
C 297	16	1.1	365	20	AAV15553
C 298	16	1.1	365	20	AAV15553
C 299	16	1.1	365	20	AAV15553
C 300	16	1.1	365	20	AAV15553
C 301	16	1.1	365	20	AAV15553
C 302	16	1.1	365	20	AAV15553
C 303	16	1.1	365	20	AAV15553
C 304	16	1.1	365	20	AAV15553
C 305	16	1.1	365	20	AAV15553
C 306	16	1.1	365	20	AAV15553
C 307	16	1.1	365	20	AAV15553
C 308	16	1.1	365	20	AAV15553
C 309	16	1.1	365	20	AAV15553
C 310	16	1.1	365	20	AAV15553
C 311	16	1.1	365	20	AAV15553
C 312	16	1.1	365	20	AAV15553
C 313	16	1.1	365	20	AAV15553
C 314	16	1.1	365	20	AAV15553
C 315	16	1.1	365	20	AAV15553
C 316	16	1.1	365	20	AAV15553
C 317	16	1.1	365	20	AAV15553
C 318	16	1.1	365	20	AAV15553
C 319	16	1.1	365	20	AAV15553
C 320	16	1.1	365	20	AAV15553
C 321	16	1.1	365	20	AAV15553
C 322	16	1.1	365	20	AAV15553
C 323	16	1.1	365	20	AAV15553
C 324	16	1.1	365	20	AAV15553
C 325	16	1.1	365	20	AAV15553
C 326	16	1.1	365	20	AAV15553
C 327	16	1.1	365	20	AAV15553
C 328	16	1.1	365	20	AAV15553
C 329	16	1.1	365	20	AAV15553
C 330	16	1.1	365	20	AAV15553
C 331	16	1.1	365	20	AAV15553
C 332	16	1.1	365	20	AAV15553
C 333	16	1.1	365	20	AAV15553
C 334	16	1.1	365	20	AAV15553
C 335	16	1.1	365	20	AAV15553
C 336	16	1.1	365	20	AAV15553
C 337	16	1.1	365	20	AAV15553
C 338	16	1.1	365	20	AAV15553
C 339	16	1.1	365	20	AAV15553
C 340	16	1.1	365	20	AAV15553
C 341	16	1.1	365	20	AAV15553
C 342	16	1.1	365	20	AAV15553
C 343	16	1.1	365	20	AAV15553
C 344	16	1.1	365	20	AAV15553
C 345	16	1.1	365	20	AAV15553
C 346	16	1.1	365	20	AAV15553
C 347	16	1.1	365	20	AAV15553
C 348	16	1.1	365	20	AAV15553
C 349	16	1.1	365	20	AAV15553
C 350	16	1.1	365	20	AAV15553
C 351	16	1.1	365	20	AAV15553
C 352	16	1.1	365	20	AAV15553
C 353	16	1.1	365	20	AAV15553
C 354	16	1.1	365	20	AAV15553
C 355	16	1.1	365	20	AAV15553
C 356	16	1.1	365	20	AAV15553
C 357	16	1.1	365	20	AAV15553
C 358	16	1.1	365	20	AAV15553
C 359	16	1.1	365	20	AAV15553
C 360	16	1.1	365	20	AAV15553

C 290	16	1.1	365	80	AW314323	AW314323 10327 MAR	363	16	1.1	427	35	C26936	C26936 C26936 Rice
C 291	16	1.1	366	31	AA279346	AA279346 zS84D07.r	C 364	16	1.1	427	38	AA201735	AA201735 LD04627.5
C 292	16	1.1	367	92	AO964435	AO964435 LTRGM17TR	C 365	16	1.1	427	109	AO612033	AO612033 HS_5085_B
C 293	16	1.1	368	59	AI779883	AI779883 EST260762	C 366	16	1.1	428	21	R00719	R00719 ye74d12.sl
C 294	16	1.1	368	61	AI822301	AI822301 LO-745T3	C 367	16	1.1	428	28	AA102196	AA102196 zN61a11.r
C 295	16	1.1	369	62	AV188724	AV188724 AV188724	C 368	16	1.1	428	50	F20388	F20388 HSPD04710 H
C 296	16	1.1	370	74	AA506854	AA506854 dn161502.s	C 369	16	1.1	430	33	AA413776	AA413776 vc57g10
C 297	16	1.1	370	79	AW241442	AW241442 xm58g08.x	C 370	16	1.1	430	38	AA766343	AA766343 oa31d06.s
C 298	16	1.1	374	33	AA448377	AA448377 zw96d06.r	C 371	16	1.1	430	91	AO847487	AO847487 LMA1FV1.1
C 299	16	1.1	374	36	C64682	C64682 C64682 Yuj1	C 372	16	1.1	431	23	H58450	H58450 yr25f12.r1
C 300	16	1.1	376	34	AA469291	AA469291 nc68d08.s	C 373	16	1.1	431	29	AA128997	AA128997 z008c11.s
C 301	16	1.1	378	50	AI683420	AI683420 LW51a08.x	C 374	16	1.1	431	41	AI005078	AI005078 ou08g01.x
C 302	16	1.1	378	69	AW123073	AW123073 UT-M-BH2.	C 375	16	1.1	431	51	AI723862	AI723862 RH121_28
C 303	16	1.1	379	93	C87903	C87903 C87903 Mous	C 376	16	1.1	431	51	AI740568	AI740568 wg16h04.x
C 304	16	1.1	380	80	AV403113	AV403113 AV403113	C 377	16	1.1	432	43	AI104427	AI104427 EST2133716
C 305	16	1.1	380	94	AO058179	AO058179 C1T-HSP-2	C 378	16	1.1	434	63	AI995292	AI995292 701681527
C 306	16	1.1	382	20	T00343	T00343 WEST01064.E	C 379	16	1.1	434	81	AA539359	AA539359 sh33h06.y
C 307	16	1.1	383	46	AI442448	AI442448 s827b01.Y	C 380	16	1.1	435	81	AA539359	AA539359 LD17457.5
C 308	16	1.1	386	50	AU062795	AU062795 AU062795	C 381	16	1.1	435	108	AO542477	AO542477 RPCT-11-3
C 309	16	1.1	387	36	AA658550	AA658550 ns91f09.s	C 382	16	1.1	436	28	AA121959	AA121959 zm24g09.r
C 310	16	1.1	388	28	AA071609	AA071609 mm71f09.r	C 383	16	1.1	436	28	AA585854	AA585854 UT-R-E0-b
C 311	16	1.1	390	34	AA489894	AA489894 bh05a10.r	C 384	16	1.1	436	51	AI724560	AI724560 RH121_5.C
C 312	16	1.1	390	43	AI186450	AI186450 qd26b08.x	C 385	16	1.1	437	53	AA396165	AA396165 vd42c01.r
C 313	16	1.1	390	50	AU063167	AU063167 AO063167	C 386	16	1.1	437	79	AW249094	AW249094 Sheared D
C 314	16	1.1	392	85	AO714320	AO714320 HS_5424.A	C 387	16	1.1	438	109	AO649657	AO649657 Sheared D
C 315	16	1.1	394	41	AI027837	AI027837 cy74f11.x	C 388	16	1.1	439	108	AO593103	AO593103 HS_5457_A
C 316	16	1.1	395	47	AI496613	AI496613 sb09a07.Y	C 389	16	1.1	440	29	AA157947	AA157947 z053b06.s
C 317	16	1.1	396	42	AI172082	AI172082 EST218077	C 390	16	1.1	440	41	AO015776	AO015776 AO015776
C 318	16	1.1	397	22	R45805	R45805 Ha207-f Adu	C 391	16	1.1	440	90	AO794062	AO794062 nbx0051L
C 319	16	1.1	397	47	AI497902	AI497902 tm97a03.x	C 392	16	1.1	441	30	AA438080	AA438080 vd22a11.s
C 320	16	1.1	397	64	AW056171	AW056171 660005B05	C 393	16	1.1	442	33	AA332744	AA332744 ma63d01.Y
C 321	16	1.1	398	64	AW050148	AW050148 UT-M-BH1-	C 394	16	1.1	442	74	AA179444	AA179444 687011C01
C 322	16	1.1	399	91	AO849499	AO849499 LMA1FV1.1	C 395	16	1.1	443	22	H10695	H10695 ym07b07.r1
C 323	16	1.1	400	94	AO035962	AO035962 C1T-HSP-2	C 396	16	1.1	443	29	AA498403	AA498403 vH43e03.r
C 324	16	1.1	401	34	AA495558	AA495558 C255 Zhou	C 397	16	1.1	446	29	AA130323	AA130323 z019f09.s
C 325	16	1.1	401	94	AO027861	AO027861 C1T-HSP-2	C 398	16	1.1	446	63	AW018112	AW018112 614067A10
C 326	16	1.1	402	40	AA978238	AA978238 ny37d07.s	C 399	16	1.1	446	106	AO399019	AO399019 mgxb0001C
C 327	16	1.1	402	46	AI428717	AI428717 vV47g08.Y	C 400	16	1.1	449	45	AI169421	AI169421 C1T-HSP-2
C 328	16	1.1	402	108	AO580428	AO580428 RPCT-11-4	C 401	16	1.1	449	94	AO037245	AO037245 EST215272
C 329	16	1.1	403	38	AA201966	AA201966 LD05033.5	C 402	16	1.1	450	32	AA366602	AA366602 EST77567
C 330	16	1.1	404	30	AA258762	AA258762 zS32a07.s	C 403	16	1.1	450	38	AA802986	AA802986 GM06718.5
C 331	16	1.1	404	44	AI278847	AI278847 q042e04.x	C 404	16	1.1	451	26	W47724	W47724 mc89e05.r1
C 332	16	1.1	404	64	AW081766	AW081766 xD70d03.x	C 405	16	1.1	451	26	AI041170	AI041170 ov69h06.x
C 333	16	1.1	405	34	AA471073	AA471073 nc70g01.r	C 406	16	1.1	452	30	AA246523	AA246523 LD05461.5
C 334	16	1.1	407	31	AA297792	AA297792 EST113341	C 407	16	1.1	452	37	AA682747	AA682747 ah52e11.s
C 335	16	1.1	408	34	AA478168	AA478168 zU40a11.r	C 408	16	1.1	453	107	AA676156	AA676156 C1TBI-EL-
C 336	16	1.1	408	45	AI372341	AI372341 a05h005.x	C 409	16	1.1	454	34	AA513452	AA513452 ng66g05.s
C 337	16	1.1	408	107	AO476927	AO476927 C1TBI-EL-	C 410	16	1.1	454	94	AO046366	AO046366 RPCT11-34
C 338	16	1.1	409	33	AA387108	AA387108 vc18e05.r	C 411	16	1.1	455	38	AA758436	AA758436 zq45b06.s
C 339	16	1.1	409	81	AA436668	AA436668 77089 MAR	C 412	16	1.1	455	85	AO687144	AO687144 nbx00073H
C 340	16	1.1	409	90	AO817479	AO817479 HS_5264.B	C 413	16	1.1	456	37	AA671324	AA671324 vn85a12.r
C 341	16	1.1	411	34	AA401343	AA401343 zU65c04.s	C 414	16	1.1	459	29	AA166052	AA166052 ms27d06.r
C 342	16	1.1	412	28	AA075567	AA075567 zM88f07.s	C 415	16	1.1	459	90	AO817039	AO817039 HS_5487.B
C 343	16	1.1	412	59	AI764289	AI764289 UT-R-Y0-a	C 416	16	1.1	459	90	AO818444	AO818444 HS_5267.A
C 344	16	1.1	413	25	N85083	N85083 J2272F Huma	C 417	16	1.1	460	33	AA448233	AA448233 zw96d07.s
C 345	16	1.1	415	21	R20456	R20456 y933e10.r1	C 418	16	1.1	461	47	AI486701	AI486701 EST2450073
C 346	16	1.1	415	33	AA419366	AA419366 zu98e01.r	C 419	16	1.1	461	47	AI512089	AI512089 LD43787.5
C 347	16	1.1	415	40	AA944644	AA944644 EST200143	C 420	16	1.1	461	106	AA409726	AA409726 HS_5095_B
C 348	16	1.1	415	41	AI013227	AI013227 EST207902	C 421	16	1.1	462	81	AA435035	AA435035 UT-R-B0P
C 349	16	1.1	415	46	AI463968	AI463968 vW68b03.Y	C 422	16	1.1	463	34	AA484865	AA484865 ne81c07.s
C 350	16	1.1	416	45	AU062828	AU062828 AU062828	C 423	16	1.1	463	34	AA484865	AA484865 vq09g08.r
C 351	16	1.1	416	55	AO679883	AO679883 HS_5457.A	C 424	16	1.1	463	74	AA173952	AA173952 f137f09.Y
C 352	16	1.1	416	80	AA058179	AA058179 C1T-HSP-2	C 425	16	1.1	465	34	AA049900	AA049900 UT-M-BH1-
C 353	16	1.1	418	69	AA123813	AA123813 UT-M-BH2.	C 426	16	1.1	466	38	AA816549	AA816549 LD02050.5
C 354	16	1.1	421	28	AA088279	AA088279 z182g07.s	C 427	16	1.1	466	39	AA869578	AA869578 vq09g08.r
C 355	16	1.1	422	28	AA108701	AA108701 mp30a11.r	C 428	16	1.1	467	41	AI023412	AI023412 cW71f08.x
C 356	16	1.1	422	59	AI775488	AI775488 EST256588	C 429	16	1.1	467	90	AO809005	AO809005 HS_5443.C
C 357	16	1.1	423	85	AO719376	AO719376 HS_5511.B	C 430	16	1.1	468	74	AA222549	AA222549 EST299360
C 358	16	1.1	424	79	AA030567	AA030567 fJ62b11.Y	C 431	16	1.1	468	23	H31711	H31711 EST106040 R
C 359	16	1.1	425	41	AI043528	AI043528 L30-531T3	C 432	16	1.1	470	39	AA825057	AA825057 ETN373.TR
C 360	16	1.1	426	25	W09299	W09299 ma63d01.r1	C 433	16	1.1	470	61	AI814897	AI814897 wk66d03.x
C 361	16	1.1	426	30	AA265998	AA265998 mz51c04.r	C 434	16	1.1	470	61	AI820357	AI820357 605092C05
C 362	16	1.1	426	79	AW297752	AW297752 UT-H-BW0-	C 435	16	1.1	470	61	AI820357	AI820357 605092C05

C 436	16	1.1	470	74	AM173074	AM173074 xj82hd04.x	C 509	16	1.1	525	81	AM400118	AM400118 707051H10
C 437	16	1.1	470	90	AO821291	AO821291 HS_4867_A	C 510	16	1.1	525	85	AO732009	AO732009 HS_5555_A
C 438	16	1.1	471	47	AI475567	AI475567 lc92d07.y	C 511	16	1.1	526	79	AM257833	AM257833 687061F09
C 439	16	1.1	471	47	AI503653	AI503653 vx73f01.x	C 512	16	1.1	527	37	AA702425	AA702425 z185d01.s
C 440	16	1.1	472	28	AA106433	AA106433 m194c04.r	C 513	16	1.1	527	49	A1629379	A1629379 fc05904.y
C 441	16	1.1	474	35	AA551857	AA551857 ng01c03.s	C 514	16	1.1	528	91	AO844005	AO844005 LMAJFV1.1
C 442	16	1.1	474	103	AO234229	AO234229 HS_2026.B	C 515	16	1.1	529	35	AA563168	AA563168 v145g09.t
C 443	16	1.1	475	42	AI140490	AI140490 ow76b01.s	C 516	16	1.1	529	92	AO949602	AO949602 Sheared D
C 444	16	1.1	475	109	AO959593	AO959593 HS_5422_B	C 517	16	1.1	531	79	AM303609	AM303609 xv21c04.x
C 445	16	1.1	477	37	AA711850	AA711850 v459c01.r	C 518	16	1.1	531	106	AO370910	AO370910 HS_5048_A
C 446	16	1.1	477	38	AA764589	AA764589 vp10e05.r	C 519	16	1.1	534	63	AI995325	AI995325 701672982
C 447	16	1.1	478	22	R22827	R22827 Yg47a08.r1	C 520	16	1.1	535	62	AI959564	AI959564 w69f03.x
C 448	16	1.1	478	92	AO943148	AO943148 Sheared D	C 521	16	1.1	535	107	AO504790	AO504790 RPCI-11-2
C 449	16	1.1	479	22	H09827	H09827 ym01a02.r1	C 522	16	1.1	537	94	AO056871	AO056871 CTR-HSP-2
C 450	16	1.1	479	37	AA725575	AA725575 a124d11.s	C 523	16	1.1	538	92	AO449536	AO449536 Sheared D
C 451	16	1.1	479	82	CNS00XON	AL095213 Arabidops	C 524	16	1.1	540	106	AO427862	AO427862 CTRB-EL-
C 452	16	1.1	480	81	AM432197	AM432197 sh70c10.y	C 525	16	1.1	541	107	AO484838	AO484838 RPCI-11-2
C 453	16	1.1	481	46	AI414601	AI414601 ma55c05.x	C 526	16	1.1	542	34	AA535891	AA535891 nE94a05.s
C 454	16	1.1	483	41	AI019726	AI019726 ua94c11.r	C 527	16	1.1	542	62	AI898937	AI898937 EST268380
C 455	16	1.1	484	24	N37830	N37830 19057 Lambda	C 528	16	1.1	542	85	AO681523	AO681523 HS_2151_A
C 456	16	1.1	486	26	W54217	W54217 md13906.r1	C 529	16	1.1	542	109	AO617200	AO617200 HS_5152_A
C 457	16	1.1	486	33	AA433010	AA433010 vq94b12.r	C 530	16	1.1	543	85	AO690409	AO690409 nbxbD0082C
C 458	16	1.1	486	59	AI777400	AI777400 EST263808	C 531	16	1.1	544	47	AI487967	AI487967 EST246289
C 459	16	1.1	487	102	AO121347	AO121347 HS_3083.B	C 532	16	1.1	544	47	AI519232	AI519232 LD38749.5
C 460	16	1.1	488	46	AI466856	AI466856 m251c04.y	C 533	16	1.1	545	44	AI297360	AI297360 LP11646.3
C 461	16	1.1	488	81	AM422401	AM422401 f162a04.y	C 534	16	1.1	546	59	AI770337	AI770337 606061B09
C 462	16	1.1	488	91	AO863528	AO863528 nbdb0021A	C 535	16	1.1	546	106	AO399673	AO399673 mgxbD0011D
C 463	16	1.1	489	34	AA516168	AA516168 n127g04.s	C 536	16	1.1	547	64	AM032841	AM032841 EST276400
C 464	16	1.1	489	62	AI899061	AI899061 EST268504	C 537	16	1.1	547	79	AM248986	AM248986 2820900.5
C 465	16	1.1	490	43	AI180866	AI180866 ub86d10.r	C 538	16	1.1	547	90	AO802947	AO802947 HS_3160_A
C 466	16	1.1	491	48	AI597150	AI597150 v192f07.x	C 539	16	1.1	549	34	AA532033	AA532033 v108e06.r
C 467	16	1.1	491	102	AO121373	AO121373 HS_3083_B	C 540	16	1.1	549	106	AO372363	AO372363 RPCI11-14
C 468	16	1.1	492	30	AA244633	AA244633 mx07g11.r	C 541	16	1.1	550	51	AI711700	AI711700 605060D09
C 469	16	1.1	492	39	C86077	C86077 C86077 Mous	C 542	16	1.1	551	105	AO311886	AO311886 RPCI11-94
C 470	16	1.1	492	35	AA597033	AA597033 v019g03.r	C 543	16	1.1	552	43	AI166902	AI166902 xy1em.est
C 471	16	1.1	493	102	AO140410	AO140410 HS_3110.B	C 544	16	1.1	553	41	AU019908	AU019908 AO019908
C 472	16	1.1	493	74	AM215916	AM215916 687046B12	C 545	16	1.1	553	62	AI913540	AI913540 w412b12.x
C 473	16	1.1	495	109	AO645364	AO645364 RPCI93-DP	C 546	16	1.1	553	107	AO504617	AO504617 RPCI-11-3
C 474	16	1.1	495	41	AU013352	AU013352 AU013352	C 547	16	1.1	554	74	AM180835	AM180835 MGAI002f
C 475	16	1.1	497	42	AI081895	AI081895 ox77b09.x	C 548	16	1.1	557	40	AA910184	AA910184 o125d04.s
C 476	16	1.1	498	45	AI333760	AI333760 q98b006.x	C 549	16	1.1	557	81	AA395134	AA395134 SH0c12.Y
C 477	16	1.1	498	63	AI967071	AI967071 456021D05	C 550	16	1.1	557	90	AO779115	AO779115 HS_3118.A
C 478	16	1.1	498	106	AO425416	AO425416 CTRB-EL-	C 551	16	1.1	558	44	AI260506	AI260506 LB04401.5
C 479	16	1.1	498	28	AA110081	AA110081 m051f10.r	C 552	16	1.1	564	61	AI820138	AI820138 605086D10
C 480	16	1.1	499	29	AA189618	AA189618 m163b09.r	C 553	16	1.1	565	26	W76805	W76805 me83e06.r1
C 481	16	1.1	500	34	AA534416	AA534416 n176c11.s	C 554	16	1.1	565	108	AO559740	AO559740 HS_5241.A
C 482	16	1.1	500	85	AO717132	AO717132 HS_2152.B	C 555	16	1.1	566	43	AI232024	AI232024 EST228712
C 483	16	1.1	501	107	AO479179	AO479179 RPCI-11-2	C 556	16	1.1	566	107	AO495364	AO495364 HS_5201.B
C 484	16	1.1	504	106	AO373151	AO373151 RPCI11-14	C 557	16	1.1	566	108	AO580448	AO580448 RPCI-11-4
C 485	16	1.1	504	109	AO612130	AO612130 HS_5111.A	C 558	16	1.1	568	79	AM231450	AM231450 687061F09
C 486	16	1.1	504	109	AO669332	AO669332 HS_5364.A	C 559	16	1.1	568	79	AM265925	AM265925 L30-2454T
C 487	16	1.1	505	106	AO396774	AO396774 mgxbD0017B	C 560	16	1.1	568	107	AO458210	AO458210 HS_5201.B
C 488	16	1.1	505	108	AO539811	AO539811 RPCI-11-3	C 561	16	1.1	571	63	AI958108	AI958108 fc90hd02.y
C 489	16	1.1	506	81	AM399389	AM399389 EST309889	C 562	16	1.1	571	102	AO157582	AO157582 nbxbD0009B
C 490	16	1.1	507	64	AM041377	AM041377 EST284241	C 563	16	1.1	572	80	AM321276	AM321276 uc29e03.y
C 491	16	1.1	509	27	W88086	W88086 m168f04.r1	C 564	16	1.1	573	94	AO016746	AO016746 CTR-HSP-2
C 492	16	1.1	511	106	AO397067	AO397067 mgxbD0010C	C 565	16	1.1	573	28	AA116947	AA116947 mg25D08.r
C 493	16	1.1	511	64	AL119787	AL119787 DRF2P761F	C 566	16	1.1	574	103	AO256656	AO256656 nbxbD0011L
C 494	16	1.1	512	85	AO703883	AO703883 HS_5501.B	C 567	16	1.1	574	64	AM078176	AM078176 fe25e12.y
C 495	16	1.1	512	107	AO455113	AO455113 HS_5049_B	C 568	16	1.1	574	102	AO158129	AO158129 nbxbD0010P
C 496	16	1.1	512	92	AO902253	AO902253 LMAJFV1.1	C 569	16	1.1	574	106	AO399361	AO399361 mgxbD0015C
C 497	16	1.1	513	109	AO668297	AO668297 HS_2125_B	C 570	16	1.1	575	38	AA792810	AA792810 vs75f08.r
C 498	16	1.1	513	74	AM217258	AM217258 EST295972	C 571	16	1.1	575	45	AU005119	AU005119 AU005119
C 499	16	1.1	515	48	AI600373	AI600373 486070E11	C 572	16	1.1	575	45	AU005120	AU005120 AU005120
C 500	16	1.1	516	30	AA220852	AA220852 mw69e07.r	C 573	16	1.1	575	81	AM412028	AM412028 uc40a05.y
C 501	16	1.1	517	46	AA986694	AA986694 uc80b07.x	C 574	16	1.1	576	34	AA485703	AA485703 pat.pK000
C 502	16	1.1	518	26	W93994	W93994 zd98b08.s1	C 575	16	1.1	576	38	AA202124	AA202124 LB02223.5
C 503	16	1.1	519	46	AI438851	AI438851 z498b08.s1	C 576	16	1.1	576	103	AO256644	AO256644 nbxbD0016L
C 504	16	1.1	520	105	AO316085	AO316085 RPCI11-10	C 577	16	1.1	576	105	AO307893	AO307893 HS_2178.B
C 505	16	1.1	520	108	AO580426	AO580426 RPCI-11.4	C 578	16	1.1	578	42	AI151842	AI151842 u146g01.y
C 506	16	1.1	522	38	AA263754	AA263754 LD07025.5	C 579	16	1.1	579	69	AM128571	AM128571 fe17e12.y
C 507	16	1.1	524	36	AA623166	AA623166 v113e05.r	C 580	16	1.1	579	36	AA660862	AA660862 00757 MCR
C 508	16	1.1	525	36	AA623166	AA623166 v113e05.r	C 581	16	1.1	582	36	AA660862	AA660862 00757 MCR

582	16	1.1	582	40	AA940672	AA940672 LD22736..5	655	16	1.1	643	51	AT745884	AT745884 605074F04
583	16	1.1	582	74	AW203846	AW203846 sE38H05..Y	656	16	1.1	647	44	AT259687	AT259687 LP03156..5
584	16	1.1	585	41	AU019901	AU019901	657	16	1.1	648	44	AT259925	AT259925 LP03449..5
585	16	1.1	586	85	AO728004	AO728004 HS..5A37..B	658	16	1.1	648	47	AT485233	AT485233 EST243537
586	16	1.1	587	64	AM058560	AM058560 sX23601..X	659	16	1.1	648	107	AO447126	AO447126 mgxB00050C
587	16	1.1	587	64	AM065185	AM065185 ST40D12..P	660	16	1.1	648	108	AO569840	AO569840 HS..5334..B
588	16	1.1	587	105	AQ328711	AQ328711 nbxD00043P	661	16	1.1	648	109	AO655457	AO655457 Sheared D
589	16	1.1	588	51	AT739863	AT739863 605072E11	662	16	1.1	650	44	AT261112	AT261112 LP05194..5
590	16	1.1	588	80	AV403084	AV403084	663	16	1.1	651	105	AO291060	AO291060 nbxB000380
591	16	1.1	589	38	AA754240	AA754240 97M00046	664	16	1.1	652	79	AW257357	AW257357 EST305494
592	16	1.1	590	34	AA502909	AA502909 n642C10..S	665	16	1.1	653	38	AA390663	AA390663 LD09636..5
593	16	1.1	592	105	AQ273070	AQ273070 nbxD00029I	666	16	1.1	654	44	AT296230	AT296230 LP10074..5
594	16	1.1	593	26	W38690	W38690 zb25D09..r1	667	16	1.1	654	44	AT260805	AT260805 LP04786..5
595	16	1.1	593	44	AT258309	AT258309 LP01463..5	668	16	1.1	654	51	AT128315	AT128315 BNLGH1104
596	16	1.1	593	62	AT1881901	AT1881901 606074H05	669	16	1.1	657	44	AT129668	AT129668 LP10649..5
597	16	1.1	593	107	AQ449447	AQ449447 mgxB00023H	670	16	1.1	658	81	AM442415	AM442415 EST311811
598	16	1.1	594	63	AT981233	AT981233 pgt..PK004	671	16	1.1	659	44	AT1260841	AT1260841 LP04828..5
599	16	1.1	594	107	AQ483757	AQ483757 RPT-11-2	672	16	1.1	661	44	AT1260733	AT1260733 LP04694..5
600	16	1.1	595	79	AM234556	AM234556 sE15f08..Y	673	16	1.1	661	49	AU056954	AU056954 AU056954
601	16	1.1	595	107	AQ448921	AQ448921 mgxB00022I	674	16	1.1	662	44	AT257897	AT257897 LP06390..5
602	16	1.1	595	107	AQ473065	AQ473065 CITBI-EI-	675	16	1.1	662	44	AT1258708	AT1258708 LP10963..5
603	16	1.1	596	92	AO369906	AO369906 LERJ24TF	676	16	1.1	664	44	AT1258708	AT1258708 LP10963..5
604	16	1.1	596	109	AO614193	AO614193 HS..5127..B	677	16	1.1	665	44	AT1260698	AT1260698 LP10066..5
605	16	1.1	598	109	AO651362	AO651362 Sheared D	678	16	1.1	665	44	AT1260698	AT1260698 LP04655..5
606	16	1.1	600	102	AQ160657	AQ160657 mgxB00006F	679	16	1.1	669	105	AA259167	AA259167 nbxB00022A
607	16	1.1	601	94	AT001134	AT001134 CIT-HSP-2	680	16	1.1	670	29	AA162891	AA162891 mg12P01..r
608	16	1.1	602	44	AT001134	AT001134	681	16	1.1	671	81	AM442274	AM442274 EST311670
609	16	1.1	604	61	AL042339	AL042339 DKEF2434P	682	16	1.1	671	29	AA162891	AA162891 mg12P01..r
610	16	1.1	605	79	AM309095	AM309095 sE93C10..Y	683	16	1.1	672	81	AV405009	AV405009 AV405009
611	16	1.1	605	102	AQ142777	AQ142777 HS..2230..B	684	16	1.1	673	29	AA141773	AA141773 CK02411..C
612	16	1.1	607	80	AM352610	AM352610 660031H06	685	16	1.1	673	44	AT1257174	AT1257174 LP05366..5
613	16	1.1	608	44	AT1259615	AT1259615 LP03071..3	686	16	1.1	681	44	AT1293421	AT1293421 LP06536..5
614	16	1.1	608	74	AM222828	AM222828 EST299639	687	16	1.1	684	44	AO794542	AO794542 nbxB000530
615	16	1.1	608	92	AO364434	AO364434 LERGM17TF	688	16	1.1	685	44	AT1293769	AT1293769 LP06982..5
616	16	1.1	609	62	AT1893981	AT1893981 mE6f04..Y	689	16	1.1	685	102	AO162297	AO162297 mgxB00112E
617	16	1.1	609	93	AO377206	AO377206 RPT-23-3	690	16	1.1	689	44	AT1293658	AT1293658 LP06843..5
618	16	1.1	610	44	AT1294776	AT1294776 LP08230..5	691	16	1.1	689	90	AO780887	AO780887 HS..3113..B
619	16	1.1	610	64	AM037140	AM037140 614018D11	692	16	1.1	692	102	AO158371	AO158371 nbxB00011A
620	16	1.1	610	80	AM317567	AM317567 s952e11..Y	693	16	1.1	696	81	AM442417	AM442417 EST311813
621	16	1.1	610	82	C15H8	AJ227340 ClonA..Int	694	16	1.1	699	105	AO272965	AO272965 nbxB00028P
622	16	1.1	610	93	AO380967	AO380967 RPT-23-2	695	16	1.1	700	80	AV403427	AV403427 AV403427
623	16	1.1	610	106	AO390451	AO390451 CITBI-EI-	696	16	1.1	701	45	AU0004949	AU0004949 AV0004949
624	16	1.1	614	102	AQ158797	AQ158797 nbxB000120	697	16	1.1	702	44	AT1296294	AT1296294 LP10151..5
625	16	1.1	614	105	AQ291530	AQ291530 nbxB00039G	698	16	1.1	702	62	AT1894230	AT1894230 md16g07..x
626	16	1.1	615	69	AM102183	AM102183 sD84d07..Y	699	16	1.1	703	45	AU006164	AU006164 AU006164
627	16	1.1	615	106	AQ397741	AQ397741 mgxB00017I	700	16	1.1	704	63	AT1971877	AT1971877 w22e609..X
628	16	1.1	616	38	AA440514	AA440514 LD14906..3	701	16	1.1	705	90	AO779611	AO779611 HS..5567..B
629	16	1.1	616	47	AT185534	AT185534 EST243655	702	16	1.1	705	103	AQ255128	AQ255128 mgxB00009B
630	16	1.1	616	81	AV406124	AV406124 AV406124	703	16	1.1	706	62	AU077635	AU077635 AU077635
631	16	1.1	616	81	AM440521	AM440521 hb91h11..x	704	16	1.1	707	93	AO381292	AO381292 RPT-23-3
632	16	1.1	618	79	AL134418	AL134418 DKEF2547F	705	16	1.1	708	80	AV400472	AV400472 AV400472
633	16	1.1	618	107	AQ448049	AQ448049 mgxB00016I	706	16	1.1	710	91	AO853579	AO853579 nbxB00024K
634	16	1.1	619	44	AT1297018	AT1297018 LP11167..5	707	16	1.1	713	44	AT1294242	AT1294242 LP07566..5
635	16	1.1	619	47	AT1520123	AT1520123 LD40127..5	708	16	1.1	713	81	AV405290	AV405290 AV405290
636	16	1.1	622	105	AQ320349	AQ320349 RPT-11-10	709	16	1.1	721	81	AV405919	AV405919 AV405919
637	16	1.1	623	85	AO690132	AO690132 nbxB00081M	710	16	1.1	723	45	AU005256	AU005256 AU005256
638	16	1.1	624	44	AT1259124	AT1259124 LP02464..5	711	16	1.1	726	44	AT1293858	AT1293858 LP07106..5
639	16	1.1	624	102	AQ160619	AQ160619 mgxB00006M	712	16	1.1	727	92	AO446189	AO446189 nbxB00062B
640	16	1.1	625	63	AT1943373	AT1943373 fE79e05..Y	713	16	1.1	729	43	C96140	C96140 C96140
641	16	1.1	626	103	AO255581	AO255581 nbxB00011F	714	16	1.1	733	105	AQ274100	AQ274100 nbxB00032D
642	16	1.1	628	60	AT1812176	AT1812176 605086D10	715	16	1.1	736	46	AT127292	AT127292 mc89e05..Y
643	16	1.1	628	82	C119E1	AU056159 AU056159	716	16	1.1	742	74	AM175277	AM175277 f134C01..Y
644	16	1.1	630	49	AU056159	AU056159	717	16	1.1	742	103	AQ249513	AQ249513 T904..T7.1
645	16	1.1	631	74	AM221797	AM221797 EST298668	718	16	1.1	744	43	C97522	C97522 C97522
646	16	1.1	632	85	AO726012	AO726012 HS..5400..B	719	16	1.1	746	45	AU002624	AU002624 AU002624
647	16	1.1	634	44	AT1259752	AT1259752 LP03239..5	720	16	1.1	752	69	AM131779	AM131779 xF34e05..x
648	16	1.1	636	60	AT1795502	AT1795502 605010D05	721	16	1.1	757	91	AO898415	AO898415 HS..3122..B
649	16	1.1	639	35	AA567395	AA567395 HL01124..5	722	16	1.1	760	105	AO273699	AO273699 nbxB00030B
650	16	1.1	639	84	B23889	B23889 F1712TF..IG	723	16	1.1	762	69	AM102119	AM102119 sD83d07..Y
651	16	1.1	640	44	AT1293753	AT1293753 LP06960..5	724	16	1.1	767	39	AM107208	AM107208 um12c08..x
652	16	1.1	640	44	AT1297823	AT1297823 LP12225..3	725	16	1.1	768	39	AA817417	AA817417 LD22693..5
653	16	1.1	640	62	AT1881368	AT1881368 606067B06	726	16	1.1	769	90	AO780466	AO780466 HS..3117..B
654	16	1.1	641	74	AM221768	AM221768 EST298579	727	16	1.1	773	42	AT115108	AT115108 u142b02..Y

728	16	1.1	774	92	AQ446967	AQ446967	801	15	1.0	179	21	D47487
729	16	1.1	785	80	AM349193	AM349193	802	15	1.0	179	70	AV247858
730	16	1.1	789	80	AV401809	AV401809	803	15	1.0	181	39	AA869973
731	16	1.1	791	107	AQ459213	AQ459213	804	15	1.0	181	80	AW311968
732	16	1.1	795	94	AQ048550	AQ048550	805	15	1.0	182	26	W78688
733	16	1.1	799	47	AQ050846	AQ050846	806	15	1.0	182	34	AA490929
734	16	1.1	802	46	AI430701	AI430701	807	15	1.0	183	32	AA351111
735	16	1.1	802	46	AV401674	AV401674	808	15	1.0	183	44	AA351111
736	16	1.1	817	80	AV401674	AV401674	809	15	1.0	184	39	AA887295
737	16	1.1	820	82	CNS01170	CNS01170	810	15	1.0	184	39	AA887295
738	16	1.1	820	105	AQ274630	AQ274630	811	15	1.0	184	62	AI794988
739	16	1.1	825	39	AA817463	AA817463	812	15	1.0	187	60	AI787690
740	16	1.1	828	102	AO161548	AO161548	813	15	1.0	189	21	R05768
741	16	1.1	829	108	AO574956	AO574956	814	15	1.0	189	51	AI716836
742	16	1.1	840	60	AI788915	AI788915	815	15	1.0	190	36	AA655332
743	16	1.1	841	91	AO895826	AO895826	816	15	1.0	191	43	AI166263
744	16	1.1	842	81	AV406086	AV406086	817	15	1.0	191	49	AV011104
745	16	1.1	854	69	AM106316	AM106316	818	15	1.0	194	46	AI409702
746	16	1.1	860	91	AO897043	AO897043	819	15	1.0	195	22	R45720
747	16	1.1	861	80	AV401275	AV401275	820	15	1.0	196	24	X93818
748	16	1.1	862	27	AA027532	AA027532	821	15	1.0	196	63	AI965149
749	16	1.1	863	85	AO749603	AO749603	822	15	1.0	199	21	D47362
750	16	1.1	884	105	AO329990	AO329990	823	15	1.0	200	37	AA717753
751	16	1.1	885	91	AO899852	AO899852	824	15	1.0	203	49	AV005457
752	16	1.1	891	90	AO787795	AO787795	825	15	1.0	204	25	W11716
753	16	1.1	895	70	AM155051	AM155051	826	15	1.0	204	27	C03138
754	16	1.1	920	105	AO330274	AO330274	827	15	1.0	205	30	AA253937
755	16	1.1	951	82	CNS0140D	CNS0140D	828	15	1.0	205	48	AI616753
756	16	1.1	951	83	AF075977	AF075977	829	15	1.0	205	59	AV105849
757	16	1.1	971	105	AO325905	AO325905	830	15	1.0	205	65	AM125212
758	16	1.1	983	63	Z98835	Z98835	831	15	1.0	208	31	AA289416
759	16	1.1	986	47	AI524108	AI524108	832	15	1.0	209	36	AA645305
760	16	1.1	1101	82	CNS000X0	CNS000X0	833	15	1.0	209	70	AA249133
761	16	1.1	1101	82	CNS001W0	CNS001W0	834	15	1.0	210	45	N55973
762	16	1.1	1101	82	CNS000D3E	CNS000D3E	835	15	1.0	210	28	AI616397
763	16	1.1	1101	82	CNS000J9E	CNS000J9E	836	15	1.0	211	41	AI030624
764	16	1.1	1118	63	AI964170	AI964170	837	15	1.0	212	37	AA667634
765	15	1.0	67	44	AI277999	AI277999	838	15	1.0	212	60	AI806234
766	15	1.0	71	48	AI581600	AI581600	839	15	1.0	213	51	AV263232
767	15	1.0	90	79	AW263826	AW263826	840	15	1.0	214	70	AI699771
768	15	1.0	91	34	AA494620	AA494620	841	15	1.0	214	40	AV232231
769	15	1.0	98	69	AM106837	AM106837	842	15	1.0	215	45	AV027148
770	15	1.0	118	42	AI152382	AI152382	843	15	1.0	216	49	AV025683
771	15	1.0	120	62	AV201286	AV201286	844	15	1.0	216	91	AO846556
772	15	1.0	121	64	AM086121	AM086121	845	15	1.0	216	31	AA329904
773	15	1.0	126	27	AA015820	AA015820	846	15	1.0	217	41	AI053241
774	15	1.0	129	28	AA082154	AA082154	847	15	1.0	217	41	AI579172
775	15	1.0	133	38	AA778275	AA778275	848	15	1.0	218	62	AI901792
776	15	1.0	135	21	T73148	T73148	849	15	1.0	219	80	AA388106
777	15	1.0	138	38	AA768260	AA768260	850	15	1.0	220	32	AA343781
778	15	1.0	144	21	R06234	R06234	851	15	1.0	221	64	AM029816
779	15	1.0	144	28	AA110201	AA110201	852	15	1.0	221	72	AV293729
780	15	1.0	144	40	AA925654	AA925654	853	15	1.0	222	24	F15263
781	15	1.0	144	86	AA347073	AA347073	854	15	1.0	222	25	N44140
782	15	1.0	145	29	AA178178	AA178178	855	15	1.0	222	84	B60593
783	15	1.0	146	20	Z19618	Z19618	856	15	1.0	222	25	N61897
784	15	1.0	148	62	AI900629	AI900629	857	15	1.0	223	30	AA237446
785	15	1.0	151	24	H88678	H88678	858	15	1.0	223	61	AA155084
786	15	1.0	151	26	W71486	W71486	859	15	1.0	223	81	AA421110
787	15	1.0	152	20	D30835	D30835	860	15	1.0	223	81	AA424209
788	15	1.0	152	26	W33117	W33117	861	15	1.0	224	51	AI713873
789	15	1.0	156	27	AA009365	AA009365	862	15	1.0	225	29	AA155235
790	15	1.0	157	63	AI947273	AI947273	863	15	1.0	225	82	HSWC17809
791	15	1.0	160	32	AA366992	AA366992	864	15	1.0	226	20	AA237446
792	15	1.0	162	37	AA687464	AA687464	865	15	1.0	226	37	C02846
793	15	1.0	165	37	AA735739	AA735739	866	15	1.0	227	32	AA337143
794	15	1.0	166	23	R85356	R85356	867	15	1.0	227	70	AV229437
795	15	1.0	170	84	B79330	B79330	868	15	1.0	228	20	T25860
796	15	1.0	172	39	AA887325	AA887325	869	15	1.0	228	60	AV143375
797	15	1.0	172	80	AW384572	AW384572	870	15	1.0	228	71	AV317440
798	15	1.0	173	51	AV059556	AV059556	871	15	1.0	228	74	AV378304
799	15	1.0	175	37	AA678304	AA678304	872	15	1.0	229	21	F05901
800	15	1.0	176	38	AA768250	AA768250	873	15	1.0	229	21	R04923

C 874	15	1.0	229	37	AA704834
C 875	15	1.0	229	61	AI832850
C 876	15	1.0	230	20	243147
C 877	15	1.0	230	24	H75753
C 878	15	1.0	232	20	M80101
C 879	15	1.0	232	32	AA341795
C 880	15	1.0	232	63	AI991714
C 881	15	1.0	232	94	AQ009027
C 882	15	1.0	233	25	W17728
C 883	15	1.0	233	48	AI590617
C 884	15	1.0	233	62	AV202548
C 885	15	1.0	233	70	AV232275
C 886	15	1.0	233	79	AW231469
C 887	15	1.0	234	34	AA512233
C 888	15	1.0	234	35	C59369
C 889	15	1.0	234	64	AM090247
C 890	15	1.0	234	80	AM381296
C 891	15	1.0	235	49	AV004176
C 892	15	1.0	236	48	AI618318
C 893	15	1.0	236	49	AV025694
C 894	15	1.0	236	80	AM378336
C 895	15	1.0	237	23	R95248
C 896	15	1.0	237	32	AA337246
C 897	15	1.0	237	50	AV044224
C 898	15	1.0	237	81	AA417416
C 899	15	1.0	237	92	AQ971909
C 900	15	1.0	238	23	H56516
C 901	15	1.0	239	69	AV214341
C 902	15	1.0	240	32	AA379565
C 903	15	1.0	240	34	AA529673
C 904	15	1.0	240	35	C40239
C 905	15	1.0	240	36	C68022
C 906	15	1.0	240	36	C68450
C 907	15	1.0	240	36	C68845
C 908	15	1.0	240	52	AV117929
C 909	15	1.0	240	59	AV186070
C 910	15	1.0	240	62	AV203565
C 911	15	1.0	241	49	AV026811
C 912	15	1.0	241	72	AV286464
C 913	15	1.0	242	25	N86470
C 914	15	1.0	242	49	AV016297
C 915	15	1.0	242	62	AI908006
C 916	15	1.0	242	73	AV370334
C 917	15	1.0	242	73	AV370407
C 918	15	1.0	242	80	AA405138
C 919	15	1.0	242	81	AA453348
C 920	15	1.0	242	82	CNS000C2
C 921	15	1.0	243	29	AA178768
C 922	15	1.0	243	59	AV071325
C 923	15	1.0	243	70	AV232576
C 924	15	1.0	243	70	AV234349
C 925	15	1.0	243	72	AV326013
C 926	15	1.0	243	72	AV326013
C 927	15	1.0	244	70	AV232555
C 928	15	1.0	244	72	AV286117
C 929	15	1.0	245	31	AA324957
C 930	15	1.0	245	33	AA389026
C 931	15	1.0	245	59	AV073378
C 932	15	1.0	245	70	AV238429
C 933	15	1.0	245	71	AV313650
C 934	15	1.0	246	49	AV016425
C 935	15	1.0	246	80	AA368833
C 936	15	1.0	247	40	AA69457
C 937	15	1.0	248	42	AA022682
C 938	15	1.0	248	80	AA361789
C 939	15	1.0	249	20	TI9295
C 940	15	1.0	249	38	AA779021
C 941	15	1.0	250	71	AV273616
C 942	15	1.0	250	71	AV313561
C 943	15	1.0	250	71	AV320817
C 944	15	1.0	250	71	AV320817
C 945	15	1.0	250	79	AA309424
C 946	15	1.0	252	36	AA636798

AA704834	zj4503.s
AI832850	at64a06.x
243147	HSCIAF081.n
H75753	yu08c08.s1
M80101	WEST00638.M
AA341795	EST47458
AI991714	w48a03.x
AQ009027	RPC111-23
W17728	mb71b01.r1
AI590617	tw24b12.x
AV202548	AV202548
AV232275	AV232275
AW231469	667065a12
AA512233	vj17a12.r
C59369	C59369 yu1
AM090247	x85f10.x
AM381296	PM1-HM029
AV004176	AV004176
AI618318	zels0081.
AV025694	AV025694
AM378336	RC2-HT021
R95248	SMOVL3CA132
AA337246	EST41948
AV044224	AV044224
AA417416	S4052.MAR
AQ971909	RPC1-23-3
H56516	yt88c05.r1
AV214341	AV214341
AA379565	EST92417
AA529673	v142a01.r
C40239	C40239 yu1
C68022	C68022 yu1
C68450	C68450 yu1
C68845	C68845 yu1
AV117929	AV117929
AV186070	AV186070
AV203565	AV203565
AV026811	AV026811
AV286464	AV286464
N86470	J8397F.Huma
AV016297	AV016297
AI908006	OV-BT161-
AV370334	AV370334
AV370407	AV370407
AA405138	UI-HF-B10
AA453348	660040F11
AI083024	ArabiDops
AA178768	mt16b01.r
AV071325	AV071325
AV232576	AV232576
AV234349	AV234349
AV326013	AV326013
AI382443	ta72s03.x
AV232555	AV232555
AV286117	AV286117
AA324957	EST27869
AA389026	mp13g11.r
AV073378	AV073378
AV238429	AV238429
AV313650	AV313650
AV016425	AV016425
AA368833	AA368833
AA69457	co81B09.s
AA022682	AA022682
AA361789	PM0-CT026
TI9295	a070038.Tes
AA779021	ab70g10.r
AI600348	486070A09
AV273616	AV273616
AV313561	AV313561
AV320817	AV320817
AA309424	sfl7e09.x
AA636798	vt16e04.r

947	15	1.0	252	39	AA835369
948	15	1.0	252	70	AV242900
949	15	1.0	253	23	D58900
950	15	1.0	253	28	AA105839
951	15	1.0	253	42	AA154536
952	15	1.0	253	71	AV273583
953	15	1.0	253	102	AQ097092
954	15	1.0	254	27	C03525
955	15	1.0	254	31	AA319242
956	15	1.0	254	34	AA869034
957	15	1.0	254	44	AU045885
958	15	1.0	254	72	AV290282
959	15	1.0	255	20	T28193
960	15	1.0	255	21	D48625
961	15	1.0	255	32	AA339088
962	15	1.0	256	33	AA448060
963	15	1.0	256	72	AV286567
964	15	1.0	257	31	AA295625
965	15	1.0	257	50	AU064301
966	15	1.0	257	50	AV032164
967	15	1.0	257	59	AV114691
968	15	1.0	257	71	AV275000
969	15	1.0	258	29	AA177646
970	15	1.0	258	60	AV123336
971	15	1.0	258	73	AV364884
972	15	1.0	259	35	C37917
973	15	1.0	259	35	C71904
974	15	1.0	260	62	AV177889
975	15	1.0	260	72	AV234591
976	15	1.0	260	107	AQ486732
977	15	1.0	261	32	AA377028
978	15	1.0	261	45	AI387317
979	15	1.0	261	46	AI454474
980	15	1.0	262	32	AA360695
981	15	1.0	262	32	AA368051
982	15	1.0	262	35	AA541430
983	15	1.0	262	70	AV230841
984	15	1.0	262	72	AV285512
985	15	1.0	262	72	AV290058
986	15	1.0	262	80	AA330998
987	15	1.0	263	40	AA808931
988	15	1.0	263	80	AA314457
989	15	1.0	264	21	D48830
990	15	1.0	264	37	AA710891
991	15	1.0	264	61	AV159043
992	15	1.0	264	82	CNS000HW
993	15	1.0	265	28	D77684
994	15	1.0	266	50	AV033518
995	15	1.0	266	60	AV150025
996	15	1.0	266	85	AO685007
997	15	1.0	266	92	AO905705
998	15	1.0	267	37	AA685264
999	15	1.0	267	49	AI664216
C1000	15	1.0	267	72	AV290097

ALIGNMENTS

RESULT	1
AA985688	286 bp
LOCUS	mus musculus
DEFINITION	mus musculus
ACCESSION	AA985688
VERSION	AA985688.1
KEYWORDS	mus musculus
SOURCE	house mouse
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 286)

AUTHORS Marra,M., Hillier L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Gaisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellhammer,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
The WashU-HM1 Mouse EST Project
Unpublished (1996)
On Jan 14, 1998 this sequence version replaced gi:1798779.
CONTACT: Marra M/Mouse EST Project
WashU-HM1 Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:928560
Seq primer: primer name ambiguous
High quality sequence stop: 273.
FEATURES
source
1. .286
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone_id="Sugano mouse embryo mewa"
/dev_stage="embryo, 14 dpc"
/lab_host="DH10B"
/note="Vector: PME18S-FL3; Site.1: DraIII (CACCCTGTG); Site.2: DraIII (CACCCTGTG); 1st strand cDNA was primed with an oligo(dT) primer (ATGTGGCTTTTCTTTTCTTTT); double-stranded cDNA was ligated to a DraIII adaptor (TGTGGCCCTACTGG), digested and cloned into distinct DraIII sites of the PME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCCTGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTCTGCTCAAAAGCTGCG and 3' end primer CGACCTGCAGCTGCAGACA."
BASE COUNT 50 a 78 c 100 g 58 t
ORIGIN
Query Match 1.3%; Score 20; DB 40; Length 286;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1321 GCCTTCGCCGTGCCACCTA 1340
|||||
Db 3 GCCTTCGCCGTGCCACCTA 22
RESULT 2
AQ444631 491 bp DNA GSS 31-MAR-1999
LOCUS GSSTC01893 Trypanosoma cruzi random genomic library Trypanosoma
DEFINITION cruzi genomic clone G8113 5', genomic survey sequence.
ACCESSION AQ444631
VERSION AQ444631.1 GI:4556095
KEYWORDS GSS.
SOURCE Trypanosoma cruzi.
ORGANISM Trypanosoma cruzi
Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma; Schizotrypanum.
REFERENCE 1 (bases 1 to 491)
AUTHORS Sanchez,D.O.
TITLE Trypanosoma cruzi random genomic sequences
JOURNAL Unpublished (1999)
COMMENT Contact: Sanchez D.O.
Instituto de Investigaciones Bioteconologicas (Univ. Nac. de Gral

San Martin)
Av. Gral Paz entre Albarcellos y Constituyentes, INTI edificio 24
CP1650) San Martin, Prov. de BS AS, Argentina
Tel: (54-11)4752-0021
Fax: (54-11)4752-9639
Email: dsanchez@int.gov.ar
Seq primer: T7
Class: Shotgun.
FEATURES
source
1. .491
/organism="Trypanosoma cruzi"
/strain="Cl-Brenner"
/db_xref="taxon:5693"
/clone="G8113"
/clone_id="Trypanosoma cruzi random genomic library"
/cell_type="epimastigote"
/note="Vector: PBS(-) (PHARMACIA)"
BASE COUNT 82 a 108 c 139 g 159 t 3 others
ORIGIN
Query Match 1.3%; Score 20; DB 107; Length 491;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1418 TCGCCAGAGAGCTCCGAG 1437
|||||
Db 443 TCGGCAGAGAGCTCCGAG 424
RESULT 3
AQ162320 713 bp DNA GSS 09-SEP-1998
LOCUS m9xb0012116r CUGI Rice Blast BAC Library Pyricularia grisea genomic
DEFINITION m9xb0012116r, genomic survey sequence.
ACCESSION AQ162320
VERSION AQ162320.1 GI:3558721
KEYWORDS GSS.
SOURCE Pyricularia grisea.
ORGANISM Pyricularia grisea
Eukaryota; Fungi; Ascomycota; amorphic Ascomycota; Pyricularia.
REFERENCE 1 (bases 1 to 713)
AUTHORS Yu,Y., Zhu,H., Boyd,C.A., Gaudette,B., Gayle,A., Kingsbury,R., Phillips,K., Sasimowski,M., Wing,R.A. and Dean,R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome
JOURNAL Unpublished (1998)
CONTACT: Dean RA
Clemson University Genomics Institute
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: GGAAACAGCTATGACCATG
Class: BAC ends
High quality sequence stop: 382.
FEATURES
source
1. .713
/organism="Pyricularia grisea"
/strain="70-15"
/db_xref="taxon:89476"
/clone="m9xb0012116r"
/clone_id="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site.1: HindIII; Site.2: HindIII;
Rice Blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library

containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25x genome coverage. High density colony filters are available upon request."

BASE COUNT
ORIGIN

103 a 257 c 167 g 186 t

Query Match 1.3%; Score 20; DB 102; Length 713;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1464 ATGCTGCATTCCTCACTGC 1483
|||||
Db 687 ATGCTGCATTCCTCACTGC 706

RESULT 4
LOCUS

AQ0904106/c 376 bp DNA GSS 30-NOV-1999
DEFINITION GSSTC05947 Trypanosoma cruzi random genomic library Trypanosoma cruzi genomic clone G36B10 5', genomic survey sequence.

ACCESSION AQ0904106
VERSION AQ0904106.1 GI:6484428
KEYWORDS GSS.
SOURCE Trypanosoma cruzi.
ORGANISM Trypanosoma cruzi.
Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma; Schizotrypanum.
1 (bases 1 to 376)
Sanchez,D.O.

REFERENCE
AUTHORS Trypanosoma cruzi random genomic sequences
TITLE Unpublished (1999)
JOURNAL Contact: Sanchez D.O.
COMMENT Instituto de Investigaciones Biotecnologicas (Univ. Nac. de Gral San Martin)
Av. Gral Paz entre Albarcellos y Constituyentes, INTI edificio 24
CP(1650) San Martin, Prov. de Bs AS, Argentina
Tel: (54-11)4752-0021
Fax: (54-11)4752-9639
Email: dsanchez@inti.gov.ar
Seq primer: T7
Class: shotgun.

FEATURES
source Location/Qualifiers
1. 376
/organism="Trypanosoma cruzi"
/strain="Cl-Brener"
/db_xref="taxon:5693"
/clone="G36B10"
/cell_type="Trypanosoma cruzi random genomic library"
/note="Vector: PBS(-) (Stratagene)"
/note="Vector: PBS(-) (Stratagene)"

BASE COUNT 60 a 104 c 108 g 102 t 2 others
ORIGIN

Query Match 1.3%; Score 19; DB 92; Length 376;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1418 TCGCGAAGAGCTCCGAA 1436
|||||
Db 175 TCGCGAAGAGCTCCGAA 157

RESULT 5
LOCUS

AA556143 392 bp mRNA EST 13-AUG-1997
DEFINITION TENF0072 T. cruzi epimastigote normalized cDNA library Trypanosoma cruzi cDNA clone 72 5', mRNA sequence.

ACCESSION AA556143
VERSION AA556143.1 GI:2327595
KEYWORDS EST.

SOURCE Trypanosoma cruzi.
ORGANISM Trypanosoma cruzi
Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma; Schizotrypanum.
1 (bases 1 to 392)
Brandao,A., Urmenyil,T., Rondinelli,E. and Degraeve,W.
Generation of EST's for the Trypanosoma cruzi genome project
Unpublished (1997)
On Sep 12, 1996 this sequence version replaced gi:1393242.
Contact: Degraeve W
Department Biochemistry & Molecular Biology
Oswaldo Cruz Institute - FIOCRUZ
Av. Brasil, 4365 - 21045-900 Rio de Janeiro-RJ Brazil
Tel: 55 21 290 7549
Fax: 55 21 590 3495
Email: wdegrave@gene.dbm.fiocruz.br
Seq primer: T7.
Class: shotgun.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source Location/Qualifiers
1. 392
/organism="Trypanosoma cruzi"
/strain="Cl-Brener"
/db_xref="taxon:5693"
/clone="72"
/cell_type="T. cruzi epimastigote normalized cDNA library"
/note="cDNA library constructed with oligo dt primed
epimastigote mRNA and cloned in pT7c318D phagemid with
modified polylinker (Pharmacia)"

BASE COUNT 113 a 108 c 110 g 58 t 3 others
ORIGIN

Query Match 1.3%; Score 19; DB 35; Length 392;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 601 CAGAGCTGATCAGCTTG 619
|||||
Db 313 CAGAGCTGATCAGCTTG 331

RESULT 6
LOCUS AQ0904862/c 401 bp DNA GSS 30-NOV-1999
DEFINITION GSSTC08025 Trypanosoma cruzi random genomic library Trypanosoma cruzi genomic clone G35N1 5', genomic survey sequence.

ACCESSION AQ0904862
VERSION AQ0904862.1 GI:6485192
KEYWORDS GSS.
SOURCE Trypanosoma cruzi.
ORGANISM Trypanosoma cruzi.
Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma; Schizotrypanum.
1 (bases 1 to 401)
Sanchez,D.O.

REFERENCE
AUTHORS Trypanosoma cruzi random genomic sequences
TITLE Unpublished (1999)
JOURNAL Contact: Sanchez D.O.
COMMENT Instituto de Investigaciones Biotecnologicas (Univ. Nac. de Gral San Martin)
Av. Gral Paz entre Albarcellos y Constituyentes, INTI edificio 24
CP(1650) San Martin, Prov. de Bs AS, Argentina
Tel: (54-11)4752-0021
Fax: (54-11)4752-9639
Email: dsanchez@inti.gov.ar
Seq primer: T7
Class: shotgun.

FEATURES
source Location/Qualifiers
1. 401
/organism="Trypanosoma cruzi"
/strain="Cl-Brener"
/db_xref="taxon:5693"
/clone="G35N1"

```
/clone_lib="Trypanosoma cruzi random genomic library"
/cell_type="epimastigote"
/note="Vector: pBS(-) (Stratagene)"

BASE COUNT      70 a   102 c   110 g   103 t   16 others
ORIGIN

Query Match      1.3%; Score 19; DB 92; Length 401;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1418 TCGGCAGAGAGACTCCGAA 1436
|||||
Db 169 TCGGCAGAGAGACTCCGAA 151

RESULT 7
AO904929/c 401 bp DNA GSS 30-NOV-1999
DEFINITION GSSTC08087 Trypanosoma cruzi random genomic library Trypanosoma
ACCESSION AO904929
VERSION AO904929.1 GI:6485259
KEYWORDS GSS.
SOURCE Trypanosoma cruzi.
ORGANISM Trypanosoma cruzi.
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

REFERENCE
AUTHORS Sanchez,D.O.
TITLE Trypanosoma cruzi random genomic sequences
JOURNAL Unpublished (1999)
COMMENT Contact: Sanchez D.O.
Instituto de Investigaciones Biotecnologicas (Univ. Nac. de Gral
San Martin)
Av. Gral Paz entre Albarellos y Constituyentes, INTI edificio 24
CP(1650) San Martin, Prov. de Bs AS, Argentina
Tel: (54-11)4752-0021
Fax: (54-11)4752-9639
Email: dsanchez@int.igov.ar
Seq primer: T7
Class: shotgun.

FEATURES
Location/Qualifiers
source 1..401
/organism="Trypanosoma cruzi"
/strain="Cl-Brener"
/db_xref="taxon:5693"
/clone="G36D3"
/clone_lib="Trypanosoma cruzi random genomic library"
/cell_type="epimastigote"
/note="Vector: pBS(-) (Stratagene)"

BASE COUNT      64 a   109 c   112 g   102 t   14 others;
ORIGIN

Query Match      1.3%; Score 19; DB 92; Length 401;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1418 TCGGCAGAGAGACTCCGAA 1436
|||||
Db 196 TCGGCAGAGAGACTCCGAA 178

RESULT 8
AM283695/c 438 bp mRNA EST 03-FEB-2000
LOCUS LG1_222_D05.g1.A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA
DEFINITION
ACCESSION AM283695
VERSION AM283695.1 GI:6673539
KEYWORDS EST.
SOURCE sorghum.
```

```
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphylliphytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
Poaceae; Sorghum.
REFERENCE
AUTHORS 1 (bases 1 to 438)
TITLE Cordonnier-Pratt,M.M., Gingle,A., Pratt,L. and Paterson,A.
JOURNAL An EST database from Sorghum
COMMENT Unpublished (1999)
On Mar 10, 1998 this sequence version replaced gi:2948910.
Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: T7
High quality sequence start: 7
High quality sequence stop: 438
POLYA=Yes.

FEATURES
Location/Qualifiers
source 1..438
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Light Grown 1 (LG1)"
/note="Organ: 10- to 14-day-old light-grown (greenhouse)
seedlings; Vector: Lambda Zap; Site 1: XhoI; Site 2:
EcoRI; The library was made from poly-A RNA in the cloning
vector lambda Zap IT. Clones to be sequenced were
prepared by mass excision."

BASE COUNT      85 a   149 c   83 g   121 t
ORIGIN

Query Match      1.3%; Score 19; DB 79; Length 438;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 511 GTGGCGGATCGATGCGG 529
|||||
Db 174 GTGGCGGATCGATGCGG 156

RESULT 9
AO662443 529 bp DNA GSS 23-JUN-1999
LOCUS HS_5455_A2_c11_17A RPCI-11 Human Male BMC Library Homo sapiens
DEFINITION genomic clone Plate=1031 Col=22 Row=E, genomic survey sequence.
ACCESSION AO662443
VERSION AO662443.1 GI:5170211
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT On Feb 19, 1999 this sequence version replaced gi:4145379.
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
```

Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCR-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
 Plate: 1031 row: E column: 22
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 529.
 Location/Qualifiers

FEATURES

source

1. 529
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="plate-1031 Col-22 Row=E"
 /clone_1lb="RPCR-11 Human Male BAC Library"
 /sex="Male"
 /note="Vector: pBACe3.6; Genomic sequence of BAC ends"

BASE COUNT 159 a 119 c 101 g 148 t 2 others
 ORIGIN

Query Match 1.3%; Score 19; DB 109; Length 529;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1435 AGACACACGCTCAGTAT 1453

Db 424 AGACACACGCTCAGTAT 442

RESULT 10 B62537 567 bp DNA GSS 21-NOV-1997
 LOCUS T22C14TF TAMU Arabidopsis thaliana genomic clone T22C14, genomic
 DEFINITION Survey sequence.
 ACCESSION B62537
 VERSION B62537.1 GI:2629299
 KEYWORDS GSS.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
 Arabidopsis.
 1 (bases 1 to 567)
 Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Linher,K.,
 Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Adams,M.D. and
 Venter,J.C.
 A BAC End Sequence Database for Identifying Minimal Overlaps in
 Arabidopsis Genomic Sequencing. Update 3
 Unpublished (1997)
 Contact: Steve Rounsley
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: rounsley@ligr.org
 Seq primer: M13-21
 Class: BAC ends
 High quality sequence stop: 567.
 Location/Qualifiers

FEATURES

source

1. 567
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone_lib="T22C14"
 /clone_1lb="TAMU"
 /sex="hermaphrodite"
 /note="Vector: Belobacti; Site_1: HindIII; Site_2:
 HindIII; Produced by Rod Wing"

BASE COUNT 100 a 193 c 141 g 133 t
 ORIGIN

Query Match 1.3%; Score 19; DB 84; Length 567;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 202 TACGAGAGCTGACAGAGA 220

Db 195 TACGAGAGCTGACAGAGA 177

RESULT 11 A1667964 572 bp mRNA EST 07-JUL-1999
 LOCUS TENG0802 T. Cruz1 epimastigote normalised cDNA library Trypanosoma
 DEFINITION cruzi cDNA clone n1218.r 5', mRNA sequence.
 ACCESSION A1667964
 VERSION A1667964.1 GI:4826336
 KEYWORDS EST.
 SOURCE Trypanosoma cruzi.
 ORGANISM Trypanosoma cruzi
 Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
 Trypanosoma; Schizotrypanum.
 1 (bases 1 to 572)
 Gonzalez Rey,E., Remisz,E., Delgado Garcia,A. and Gonzalez,A.
 Characterization of ESTs from Trypanosoma cruzi epimastigotes
 Unpublished (1998)
 On May 18, 1998 this sequence version replaced g1:3137105.
 Contact: Delgado Alberto
 Departamento de Biologia Molecular, lab 303
 Instituto de Parasitologia y Biomedicina
 Consejo Superior de Investigaciones Cientificas C/ Ventanilla No
 11, E-18001, Granada, Spain
 Tel: 34 958 805058
 Fax: 34 958 203323
 Similar to 40S ribosomal protein S25, sp|P46301|RS25_LYCES with
 Blastx.
 Seq primer: T7
 High quality sequence stop: 572.
 Location/Qualifiers

FEATURES

source

1. 572
 /organism="Trypanosoma cruzi"
 /strain="Cl - Brener"
 /db_xref="taxon:5693"
 /clone_lib="n1218.r"
 /clone_1lb="T. Cruz1 epimastigote normalised cDNA library"
 /cell_type="epimastigote"
 /note="Site_1: EcoRI; Site_2: NotI; cDNA library
 constructed with oligo dt primed epimastigote mRNA and
 cloned in pTt318D phagemid with modified polylinker"

BASE COUNT 155 a 128 c 145 g 144 t
 ORIGIN

Query Match 1.3%; Score 19; DB 50; Length 572;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 601 CAGAGCTGATCAAGCTTG 619

Db 318 CAGAGCTGATCAAGCTTG 336

RESULT 12 A1667889 587 bp mRNA EST 07-JUL-1999
 LOCUS TENG0838 T. Cruz1 epimastigote normalised cDNA library Trypanosoma
 DEFINITION cruzi cDNA clone n407.r 5', mRNA sequence.
 ACCESSION A1667889
 VERSION A1667889.1 GI:4826261
 KEYWORDS EST.

SOURCE
Trypanosoma cruzi.
Trypanosoma cruzi
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma; Schizotrypanum.

REFERENCE
1 (bases 1 to 587)
Gonzalez Rey,E., Remisz,E., Delgado Garcia,A. and Gonzalez,A.
Characterization of ESTs from Trypanosoma cruzi epimastigotes
Unpublished (1998)

JOURNAL
On May 18, 1998 this sequence version replaced gi:3137030.
Contact: Delgado Alberto
Departamento de Biologia Molecular, Lab 303
Instituto de Parasitologia y Biomedicina
Consejo Superior de Investigaciones Cientificas C/ Ventanilla No
11, E-18001, Granada, Spain
Tel: 34 958 805058
Fax: 34 958 203323

COMMENT
Similar to 405 ribosomal protein S25 sp|P46301|RS25.LYCES with
Blastx.
Seq primer: T7
High quality sequence stop: 587.
Location/Qualifiers
1..587
/organism="Trypanosoma cruzi"
/strain="CI - Brenner"
/db_xref="taxon:5693"
/clone="n407.r"
/clone_lib="T. Cruzi epimastigote normalised cDNA library"
/cell_type="epimastigote"
/note="Site_1: EcoRI; Site_2: NotI; cDNA library
constructed with oligo dt primed epimastigote mRNA and
cloned in pTc318D phagemid with modified polylinker"
BASE COUNT
161 a 133 c 150 g 143 t
ORIGIN

Query Match
Best Local Similarity 1.3%; Score 19; DB 50; Length 587;
100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 601 CAGAGCTGATCAGCTTG 619
|||||
Db 318 CAGAGCTGATCAGCTTG 336

RESULT 13
A1527261/c 235 bp mRNA EST 18-MAR-1999
LOCUS u150f05.x1 Sugano mouse liver mila Mus musculus cDNA clone
DEFINITION IMAGE:1923393 3' similar to TR:008677 008677 HMM PREKININOGEN
PRECUSOR. ; mRNA sequence.
ACCESSION A1527261
VERSION A1527261.1 GI:4441396
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 235)
Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Rowers,Y.,
Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schuck,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
On Jun 5, 1998 this sequence version replaced gi:187218.
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:979685
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: custom primer used
High quality sequence stop: 1.
Location/Qualifiers
1..235
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1923393"
/clone_lib="Sugano mouse liver mila"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
(CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCCTACTGG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTTCGCTCTTAAAGCTGG and 3' end
primer CGACCTGCACCTCGACGACCA."

BASE COUNT
65 a 49 c 55 g 66 t
ORIGIN

Query Match
Best Local Similarity 1.2%; Score 18; DB 47; Length 235;
100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 GCTGGAGATCCCAACA 119
|||||
Db 60 GCTGGAGATCCCAACA 43

RESULT 14
C32880/c 300 bp mRNA EST 18-OCT-1999
LOCUS C32880 yuji Kohara unpublished cDNA:Strain N2 hermaphrodite embryo
DEFINITION Caenorhabditis elegans cDNA clone yk336d9 3', mRNA sequence.
ACCESSION C32880
VERSION C32880.1 GI:2364676
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditiia; Rhabditida;
Rhabditina; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 300)
Kohara,Y., Mochizuki,T., Tabara,H., Watanabe,H., Sugimoto,A.,
Sano,M., Miyata,A. and Nishigaki,A.
Expression map of the C.elegans genome
Unpublished (1996)
On Sep 12, 1996 this sequence version replaced gi:1404978.
Contact: Yuji Kohara
Gene Library Lab
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: yk@haraelab.nig.ac.jp.
Location/Qualifiers
1..300
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk336d9"

BASE COUNT 55 a 89 c 55 g 100 t 1 others

ORIGIN

/clone_1lb="Yuji Kohara unpublished cDNA:Strain N2
hermaphrodite embryo"
/sex="hermaphrodite"
/dev_stage="embryo"

Query Match 1.2%: Score 18; DB 35; Length 300;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 202 TACGAGAGCTGAGAGAG 219
|||||

Db 148 TACGAGAGCTGAGAGAG 131

RESULT 15
AV175791 300 bp mRNA EST 21-JUL-1999
LOCUS AV175791 Yuji Kohara unpublished cDNA:Strain N2 hermaphrodite
DEFINITION embryo Caenorhabditis elegans cDNA clone yk46396 3', mRNA sequence.
ACCESSION AV175791
VERSION AV175791.1 GI:5555692
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
Rhabditina; Rhabditoidea; Rhabditidae; Pelodolerinae; Caenorhabditis.
1 (bases 1 to 300)
Kohara, Y., Shin-I, T., Thierly-Mieg, J., Thierly-Mieg, D., Mitsuki, H.,
Nishigaki, A., Motobashi, T., Zeng, Q., Watanabe, H., Sugimoto, A.,
Sano, M., Miyata, A., Mitani, Y., Iida, K., Uesugi, H., Sugiyama, Y. and
Nomoto, H.

REFERENCE
AUTHORS
TITLE Expressed genes in C. elegans
JOURNAL Unpublished (1999)
COMMENT On May 9, 1996 this sequence version replaced gi:1132756.
Contact: Yuji Kohara
Gene Library Lab
National Institute of Genetics
Yata 1111 Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
1. 300

FEATURES
SOURCE
1. 300
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk46396"
/clone_1lb="Yuji Kohara unpublished cDNA:Strain N2
hermaphrodite embryo"
/sex="hermaphrodite"
/dev_stage="embryo"

BASE COUNT 56 a 89 c 54 g 101 t

ORIGIN

Query Match 1.2%: Score 18; DB 62; Length 300;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 202 TACGAGAGCTGAGAGAG 219
|||||

Db 147 TACGAGAGCTGAGAGAG 130

Search completed: May 10, 2000, 23:09:48
Job time: 3939 sec

gb_htg6:AC018480	-	8.00	94.72	3.2e+03	75162	AC018480 Drosophila melanogaster	gb_htg7:AC007415	-	8.00	90.69	5.4e+03	139133	AC007415 Drosophila melanogaster
gb_in2:AC005111	-	8.00	94.69	3.2e+03	75551	AC005111 Drosophila melanogaster	gb_htg7:AC018957	-	8.00	90.66	5.4e+03	139795	AC018957 Homo sapiens chr
gb_htg6:AC020238	+	8.00	94.58	3.3e+03	76803	AC020238 Drosophila melanogaster	gb_p11:AP000815	+	8.00	90.54	5.5e+03	142418	AP000815 Oryza sativa gen
gb_p11:AP000330	-	8.00	94.14	3.4e+03	82116	AP000330 Homo sapiens genomic	gb_pr3:HS64411	-	8.00	90.40	5.6e+03	145593	AL035665 Human DNA sequen
gb_htg5:AC0018150	-	8.00	94.03	3.5e+03	83551	AC018150 Drosophila melanogaster	gb_htg1:AC0136110	-	8.00	90.36	5.6e+03	146425	AL136110 Homo sapiens chr
gb_in2:AC004433	-	8.00	93.85	3.6e+03	85862	AC004433 Drosophila melanogaster	gb_pr2:HS50J22	+	8.00	90.35	5.6e+03	147646	Z84484 Human DNA sequence
gb_in2:AC004716	-	8.00	93.81	3.6e+03	86391	AC004716 Drosophila melanogaster	gb_htg11:AC022636	+	8.00	90.30	5.6e+03	147860	AC022636 Homo sapiens chr
gb_htg5:AC015732	+	8.00	93.80	3.6e+03	86464	AC015732 Homo sapiens clone H	gb_htg12:AC022907	-	8.00	90.20	5.7e+03	149551	AC022907 Homo sapiens chr
gb_htg3:AC008729	+	8.00	93.73	3.6e+03	87442	AC008729 Homo sapiens chromos	gb_p11:AP0001575	+	8.00	90.14	5.7e+03	151470	AC011575 Homo sapiens chr
gb_htg12:AC010691	+	8.00	93.60	3.7e+03	89203	AC010691 Drosophila melanogaster	gb_p11:AP000237	+	8.00	90.09	5.8e+03	152702	AP000237 Homo sapiens chr
gb_p12:ATAC011020	+	8.00	93.43	3.8e+03	91530	AC011020 Arabidopsis thaliana	gb_in1:PFWAL3P2	+	8.00	90.07	5.8e+03	153098	AL034558 Plasmodium falci
gb_pr4:AC004596	-	8.00	93.35	3.8e+03	92638	AC004596 Homo sapiens chromos	gb_htg8:AC022823	+	8.00	90.06	5.8e+03	153379	AC022823 Homo sapiens chr
gb_in2:AC005148	-	8.00	93.17	3.9e+03	95262	AC005148 Drosophila melanogaster	gb_htg12:AC008979	+	8.00	90.05	5.8e+03	153433	AC008979 Homo sapiens chr
gb_p12:AC004044	-	8.00	92.95	4.0e+03	98527	AC004044 Arabidopsis thaliana	gb_pr4:AF107258	+	8.00	90.02	5.8e+03	154311	AF107258 Homo sapiens chr
gb_htg3:AC009911	+	8.00	92.87	4.0e+03	99707	AC009911 Drosophila melanogaster	gb_htg11:AC008231	+	8.00	90.00	5.9e+03	154729	AC008099 Drosophila melanogaster
gb_htg3:AC009911	+	8.00	92.87	4.0e+03	99707	AC009911 Drosophila melanogaster	gb_htg11:AC008231	+	8.00	90.00	5.9e+03	154729	AC008099 Drosophila melanogaster
gb_p11:AP000092	-	8.00	92.85	4.1e+03	100000	AP000092 Homo sapiens genomic	gb_htg12:AC017063	+	8.00	89.93	5.9e+03	156327	AL137063 Homo sapiens chr
gb_p11:AP000196	-	8.00	92.85	4.1e+03	100000	AP000196 Homo sapiens genomic	gb_htg12:AC012389	+	8.00	89.92	5.9e+03	156725	AC012389 Drosophila melanogaster
gb_p11:AP000216	-	8.00	92.85	4.1e+03	100000	AP000216 Homo sapiens genomic	gb_htg12:AC012389	+	8.00	89.92	5.9e+03	156725	AC012389 Drosophila melanogaster
gb_p11:AP000517	+	8.00	92.85	4.1e+03	100000	AP000517 Homo sapiens genomic	gb_htg3:HSJ1059L7	+	8.00	89.89	5.9e+03	157347	AL121913 Homo sapiens chr
gb_htg8:AC008362	-	8.00	92.82	4.1e+03	100477	AC008362 Drosophila melanogaster	gb_htg3:AC008331	+	8.00	89.86	6.0e+03	158018	AC008331 Drosophila melanogaster
gb_htg3:AC008227	+	8.00	92.75	4.1e+03	101579	AC008227 Drosophila melanogaster	gb_htg11:AC008337	+	8.00	89.83	6.0e+03	158802	AC008337 Homo sapiens chr
gb_htg5:AC009341	+	8.00	92.72	4.1e+03	102044	AC009341 Drosophila melanogaster	gb_htg11:AC009367	+	8.00	89.76	6.0e+03	160399	AC009367 Homo sapiens chr
gb_htg2:HSJ741H3	+	8.00	92.54	4.2e+03	104901	AL132773 Homo sapiens chrom	gb_htg11:AC021773	+	8.00	89.69	6.1e+03	162165	AC021773 Homo sapiens chr
gb_htg1:HSJ214H10_4	+	8.00	92.48	4.3e+03	105892	Continuation (5 of 5) of HSN	gb_p11:AB023054	+	8.00	89.68	6.1e+03	162383	AB023054 Homo sapiens chr
gb_htg11:AC007578	-	8.00	92.35	4.3e+03	107932	AC007578 Drosophila melanogaster	gb_p11:AB023054	+	8.00	89.68	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:AC013324	-	8.00	92.32	4.3e+03	108569	AC013324 Homo sapiens chrom	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb_p11:AB023054	+	8.00	89.67	6.1e+03	162383	AB023054 Homo sapiens chr
gb_p12:HS112K5	-	8.00	92.24	4.4e+03	108569	Continuation (3 of 4) of HSN	gb						

gb_bt92:HSJ509119	-	8.00	88.57	7.0e+03	192522	AL121834 Homo sapiens chrom	gb_sts:G25604	-	7.00	110.93	399.70	336	G25604 human STS EST302218,
gb_pr4:AC0081173	+	8.00	88.52	7.1e+03	193893	AC0081173 Homo sapiens clone	gb_p11:NS228033	+	7.00	110.91	400.70	337	AJ228033 Nicotiana sylvestris,
gb_bt93:AC009502	-	8.00	88.52	7.1e+03	194133	AC009502 Homo sapiens clone	gb_p11:NT2227999	+	7.00	110.87	402.69	339	AJ227999 Nicotiana tabacum D
gb_bt97:AC011492	-	8.00	88.48	7.1e+03	195320	AC011492 Homo sapiens clone	gb_sts:G25497	+	7.00	110.75	408.66	345	G25497 human STS EST164465,
gb_bt97:AC015384	-	8.00	88.40	7.2e+03	195766	AC015384 Mus musculus clone	gb_ov:DRU57657	+	7.00	110.70	411.64	348	U57657 Danio rerio lipoprotein
gb_pr3:CNS01DDH	-	8.00	88.31	7.3e+03	200368	AL133244 RAC sequence from t	gb_v1:REXML1	+	7.00	110.60	414.61	351	AJ13244 Xenotropic murine leu
gb_bt97:AC016575	-	8.00	88.19	7.4e+03	200318	AC016575 Homo sapiens chrom	gb_pr2:HS233355	+	7.00	110.46	424.49	361	AJ233355 Homo sapiens deriva
gb_bt99:AC011113	-	8.00	88.14	7.4e+03	205649	AC011113 Homo sapiens clone	gb_in2:AF000784	+	7.00	110.46	424.49	361	AJ233355 Homo sapiens deriva
gb_bt99:AC009856	-	8.00	88.14	7.4e+03	205649	AC009856 Homo sapiens clone	gb_in2:AF000784	+	7.00	110.40	427.45	364	AF000784 Theatops erythrocep
gb_bt99:AC010864	+	8.00	88.13	7.4e+03	205670	AC010864 Homo sapiens clone	gb_pr2:HS233356	+	7.00	110.39	428.43	365	AF000784 Homo sapiens parlia
gb_bt94:AC012108	-	8.00	88.09	7.5e+03	207284	AC012108 Homo sapiens clone	gb_in2:AF000779	+	7.00	110.39	428.43	365	AF000779 Scutigera coleoptra
gb_bt94:BSUR0005	-	8.00	88.05	7.5e+03	208430	AC012108 Homo sapiens clone	gb_p11:HSMD2A12	+	7.00	110.31	432.37	369	X72423 H.sapiens MRNA for re
gb_bt97:AC009167	-	8.00	87.98	7.6e+03	210824	Z99108 Bacillus subtilis con	gb_p11:HSMD2A12	+	7.00	110.19	439.23	376	X72423 H.sapiens MRNA for re
gb_bt97:AC022061	-	8.00	87.86	7.6e+03	214521	AC022061 Mus musculus clone	gb_ov:PRBLEX1	+	7.00	110.17	440.21	377	F53327 F.rubripes butyrphol
gb_bt910:AC008513	-	8.00	87.82	7.7e+03	216105	AC022061 Mus musculus clone	gb_pr4:AF212944	+	7.00	110.05	447.06	384	AF212944 Galagoides zanzibar
gb_bt93:HSJ256022	-	8.00	87.81	7.7e+03	216215	AL080513 Homo sapiens chrom	gb_sts:G00743	+	7.00	110.04	448.03	385	G00743 fruit fly STS Dm0633
gb_bt92:AL139344	+	8.00	87.67	7.9e+03	220897	AL080513 Homo sapiens chrom	gb_pr4:AF212944	+	7.00	110.04	448.03	385	AF212944 Otollemur classica
gb_bt95:AC018175	-	8.00	87.60	8.0e+03	223466	AL139344 Homo sapiens chrom	gb_sts:HSF248H04	+	7.00	109.89	456.80	394	AF01086 H.sapiens flavo-sort
gb_bt95:AC011639	-	8.00	87.58	8.0e+03	223988	AC018175 Homo sapiens clone	gb_sts:G51943	+	7.00	109.56	476.17	414	G51943 SHGC-84574 Human Hom
gb_bt99:AC010147	+	8.00	87.47	8.1e+03	227923	AC011639 Homo sapiens clone	gb_p11:MC299243	+	7.00	109.55	477.13	415	G29243 Medicago coronata ext
gb_bt911:AC023469	+	8.00	87.45	8.1e+03	228606	AC010147 Homo sapiens clone	gb_pr2:HS2018926	+	7.00	109.48	480.98	419	Y18926 Homo sapiens wild typ
gb_bt92:AL139236	+	8.00	87.35	8.2e+03	231982	AL139236 Homo sapiens chrom	gb_ba1:DEINORPEAT	+	7.00	109.37	487.71	426	M94966 Deinococcus radiodura
gb_bt913:AC012154	+	8.00	86.92	8.4e+03	247861	AC012391 Homo sapiens clone	gb_sts:G23480	+	7.00	109.37	487.71	426	G23480 human STS SHGC-33762,
gb_bt910:AC021035	-	8.00	86.92	8.4e+03	247861	AC012391 Homo sapiens clone	gb_sts:G23480	+	7.00	109.37	487.71	426	G23480 human STS SHGC-33762,
gb_bt910:AC021035	-	8.00	86.73	8.9e+03	255032	AC012035 Homo sapiens clone	gb_ba1:AB007299	+	7.00	109.19	499.21	438	AB007299 Unidentified archae
gb_bt97:AC009196	-	8.00	85.98	9.8e+03	286205	AC009196 Homo sapiens chrom	gb_ba1:AB007301	+	7.00	109.16	501.12	440	AB007301 Unidentified archae
gb_bt911:AC020967	-	8.00	85.33	1.1e+04	316203	AC020967 Mus musculus clone	gb_ba1:AB007302	+	7.00	109.16	501.12	440	AB007302 Unidentified archae
gb_bt911:AC021154	-	7.00	126.72	52.71	30	X66573 R.rattus Hnf-1 gene (last	gb_sts:G28440	+	7.00	109.12	503.98	443	G28440 human STS SHGC-36315,
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 7 from patent US	gb_sts:G28440	+	7.00	109.12	503.98	443	G28440 human STS SHGC-36315,
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ba2:UCU77575	+	7.00	108.76	520.56	471	U52220 Danio rerio Mel-1b me
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 7 from patent US	gb_ov:MS285973	+	7.00	108.72	530.56	471	MS285973 Musa sp. DNA for sequ
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_pr2:HSN1293R	+	7.00	108.65	535.28	476	HSN1293R Medicago rigida
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30	55.64	32	I23373 Sequence 8 from patent US	gb_ov:AF094639	+	7.00	108.51	544.70	486	AF094639 Passer domesticus C
gb_bt911:AC021154	-	7.00	126.30										

gb_pl2:CONS01B9M	7.00	107.82	595.01	540	AL113954	Botrytis cinerea strain	AL113954	Botrytis cinerea strain	gb_pr4:HSTHVGL6	7.00	105.34	818.63	790	AF105686	Homo sapiens thyrong
gb_in1:PAU078255	7.00	107.82	595.01	540	U07825	Parechocytis angulosus his	U07825	Parechocytis angulosus his	gb_ba1:MLTPTAT	7.00	105.22	830.78	804	X60720	M.lepreae tRNA-Pro gene
gb_pl1:ABO07313	7.00	107.74	601.47	547	AJ246965	Aschochyta rabiei micro	AJ246965	Aschochyta rabiei micro	gb_ba1:ABO07313	7.00	105.22	834.24	808	AB007313	Unidentified archae
gb_to:RATSVFG3	7.00	107.69	605.16	551	M1905	Rat androgen-responsive	M1905	Rat androgen-responsive	gb_pl1:ABO115948	7.00	105.10	843.75	819	AB0115948	Flammulina velutipe
gb_ba1:BIFIGSA	7.00	107.68	606.08	552	L35968	Bifidobacterium choerinu	L35968	Bifidobacterium choerinu	gb_ro:ABO2657S01	7.00	104.94	861.86	840	AB026570	Mus musculus genome f
gb_sts:G51434	7.00	107.65	608.84	555	G51434	SHGC-80923 Human Homo sa	G51434	SHGC-80923 Human Homo sa	gb_sts:GNS01ILA	7.00	104.91	864.44	843	AL115871	Anopheles gambiae S
gb_pl2:AF08479	7.00	107.51	619.86	567	E08479	gDNA encoding phospholip	E08479	gDNA encoding phospholip	gb_in2:AF201326	7.00	104.80	877.32	858	AF021326	Nemertinoidea elong
gb_pat:AF09568	7.00	107.49	620.78	568	AF09568	Gossypium hirsutum H+	AF09568	Gossypium hirsutum H+	gb_ba2:AF201355	7.00	104.76	879.89	861	AL12338	tfd-2.4-b-monoxyg
gb_pr3:HUMZD52G03	7.00	107.49	620.78	568	AF086316	Homo sapiens full leng	AF086316	Homo sapiens full leng	gb_ba2:AF201362	7.00	104.76	881.61	863	AF201355	Uncultured crenarch
gb_pr2:AF052954	7.00	107.46	623.53	571	L16258	Myriostoma coliforme	L16258	Myriostoma coliforme	gb_ba2:AF201362	7.00	104.76	881.61	863	AF201360	Uncultured crenarch
gb_sts:HUMUT465	7.00	107.45	624.53	572	G63067	SHGC-140661 Human Homo s	G63067	SHGC-140661 Human Homo s	gb_ba2:AF201362	7.00	104.76	881.61	863	AF201362	Uncultured crenarch
gb_pl2:CONS01AGF	7.00	107.40	628.10	576	AL112903	Botrytis cinerea strain	AL112903	Botrytis cinerea strain	gb_ba2:AF201363	7.00	104.76	881.61	863	K00011	Murine leukemia virus
gb_to:WSPRPM1	7.00	107.28	638.15	587	M1900	Mouse 15-kDa proline-rich	M1900	Mouse 15-kDa proline-rich	gb_v1:MLVENYX	7.00	104.76	881.61	863	AF201359	Uncultured crenarch
gb_sts:G09498	7.00	107.17	647.25	597	G09498	human STS CHLC.GC95C11.F	G09498	human STS CHLC.GC95C11.F	gb_ba2:AF201369	7.00	104.75	882.46	864	AF201365	Uncultured crenarch
gb_pr2:Y08126	7.00	107.14	649.98	600	Y08126	H.sapiens PRX gene, exon	Y08126	H.sapiens PRX gene, exon	gb_ba2:AF201368	7.00	104.75	882.46	864	AF201368	Uncultured crenarch
gb_ba2:AF052954	7.00	106.98	662.67	614	AF052954	Uncultured marine arch	AF052954	Uncultured marine arch	gb_ba2:AF201358	7.00	104.74	883.32	865	AF201356	Uncultured crenarch
gb_pl1:BMU59888	7.00	106.81	678.02	631	U95888	Baphia madagascariensis	U95888	Baphia madagascariensis	gb_ba2:AF201358	7.00	104.74	883.32	865	AF201358	Uncultured crenarch
gb_sts:G06348	7.00	106.73	684.32	638	G06348	human STS WI-7008	G06348	human STS WI-7008	gb_ba2:AF201366	7.00	104.74	883.32	865	AF201366	Uncultured crenarch
gb_sts:G41444	7.00	106.71	686.12	640	G41444	Z3399 Zebrafish AB Danio	G41444	Z3399 Zebrafish AB Danio	gb_ba2:AF201367	7.00	104.74	884.18	866	AF201367	Uncultured crenarch
gb_pr3:U086228	7.00	106.62	694.20	649	U08228	Homo sapiens olfactory r	U08228	Homo sapiens olfactory r	gb_pr2:HSAN2MLR12	7.00	104.60	899.56	884	U06989	Human alpha-2-macrog
gb_pr3:U086228	7.00	106.62	694.20	649	U08228	Homo sapiens olfactory r	U08228	Homo sapiens olfactory r	gb_sts:GNS01FRM	7.00	104.60	899.56	884	AL142211	Anopheles gambiae S
gb_pr2:HSAN018927	7.00	106.55	700.48	656	Y18927	Homo sapiens ALL-1/MLL/R	Y18927	Homo sapiens ALL-1/MLL/R	gb_pr1:HUMZ1NFCIN	7.00	104.59	900.41	885	L14864	Human zinc finger MRN
gb_pl2:CONS0195A	7.00	106.51	704.06	660	AL112066	Botrytis cinerea strain	AL112066	Botrytis cinerea strain	gb_ba1:ASPL16SSA1	7.00	104.58	902.12	887	Y08984	Archaea sp. 16S rRNA
gb_pl2:CONS01AAW	7.00	106.51	704.06	660	AL112488	Botrytis cinerea strain	AL112488	Botrytis cinerea strain	gb_ba1:ASPL16SSA2	7.00	104.58	902.12	887	Y08985	Archaea sp. 16S rRNA
gb_pl2:CONS01AQA	7.00	106.51	704.06	660	AL113274	Botrytis cinerea strain	AL113274	Botrytis cinerea strain	gb_ba1:ASPFPSB5	7.00	104.56	903.83	889	X96693	Archaea sp. 16S rRNA
gb_pl2:CONS01B7A	7.00	106.51	704.06	660	AL113870	Botrytis cinerea strain	AL113870	Botrytis cinerea strain	gb_sts:GNS01GV8	7.00	104.56	903.83	889	AL113637	Anopheles gambiae S
gb_pl2:CONS01BCU	7.00	106.51	704.06	660	AL114070	Botrytis cinerea strain	AL114070	Botrytis cinerea strain	gb_ba1:ASPFPSB1	7.00	104.54	907.23	893	X96688	Archaea sp. 16S rRNA
gb_pl2:CONS01BF7	7.00	106.51	704.06	660	AL114155	Botrytis cinerea strain	AL114155	Botrytis cinerea strain	gb_ba1:ASPFPSB10	7.00	104.54	907.23	893	X96695	Archaea sp. 16S rRNA
gb_pl2:CONS01C1M	7.00	106.51	704.06	660	AL114961	Botrytis cinerea strain	AL114961	Botrytis cinerea strain	gb_ba1:ASPFPSB11	7.00	104.54	907.23	893	X96696	Archaea sp. 16S rRNA
gb_pl2:CONS01D2M	7.00	106.51	704.06	660	AL116294	Botrytis cinerea strain	AL116294	Botrytis cinerea strain	gb_ba1:ASPFPSB2	7.00	104.54	907.23	893	X96689	Archaea sp. 16S rRNA
gb_pl2:CONS01D9C	7.00	106.51	704.06	660	AL116536	Botrytis cinerea strain	AL116536	Botrytis cinerea strain	gb_ba1:ASPFPSB3	7.00	104.54	907.23	893	X96690	Archaea sp. 16S rRNA
gb_pl1:HSAN224523	7.00	106.44	710.31	667	AJ224523	Homo sapiens TRK gene	AJ224523	Homo sapiens TRK gene	gb_ba1:ASPFPSB4	7.00	104.54	907.23	893	X96691	Archaea sp. 16S rRNA
gb_v1:AF025478	7.00	106.43	711.21	668	AF025478	La Crosse virus nucle	AF025478	La Crosse virus nucle	gb_ba1:ASPFPSB5	7.00	104.54	907.23	893	X96692	Archaea sp. 16S rRNA
gb_v1:AF025479	7.00	106.38	715.67	673	AF025479	La Crosse virus protot	AF025479	La Crosse virus protot	gb_ba1:ASPFPSB7	7.00	104.54	907.23	893	X96694	Archaea sp. 16S rRNA
gb_m:AF012929	7.00	106.38	715.67	673	AF012929	Bubalus arnee bubalis	AF012929	Bubalus arnee bubalis	gb_ba1:CREP5919	7.00	104.54	907.23	893	AJ006919	Crenarchaeote 16S r
gb_to:AF051457	7.00	106.28	725.46	684	AF051457	Rattus norvegicus beta	AF051457	Rattus norvegicus beta	gb_ba1:CREP5920	7.00	104.53	908.09	894	AJ006920	Crenarchaeote 16S r
gb_pr2:HSAN235332	7.00	106.25	728.13	687	AJ235332	Homo sapiens partial	AJ235332	Homo sapiens partial	gb_ba1:CREP5921	7.00	104.53	908.09	894	AJ006921	Crenarchaeote 16S r
gb_sts:MDU96812	7.00	106.15	727.93	687	U96812	Malus domestica DNA mark	U96812	Malus domestica DNA mark	gb_ba1:CREP6922	7.00	104.51	909.79	896	AJ006922	Crenarchaeote 16S r
gb_m:RNCPTPI	7.00	106.15	728.13	698	X05722	pig mRNA for prepropar	X05722	pig mRNA for prepropar	gb_ov:U01301	7.00	104.51	909.79	896	U01301	Sparus aurata growth
gb_pr3:HSMCUCIN5B1	7.00	106.13	739.67	700	U78552	Homo sapiens mucin (MUC5	U78552	Homo sapiens mucin (MUC5	gb_ov:AF064828	7.00	104.48	913.19	900	AF064828	Danio rario putativ
gb_pr1:HSGLGCV356	7.00	106.08	744.10	705	X74288	H.sapiens germeline IGLV3	X74288	H.sapiens germeline IGLV3	gb_pr4:AF135098	7.00	104.48	913.19	900	AF135098	Homo sapiens NR-RN
gb_pl2:CONS01C1C	7.00	106.04	748.52	710	AL115354	Botrytis cinerea strain	AL115354	Botrytis cinerea strain	gb_ba1:AGLRNNA8	7.00	104.48	914.05	901	M88075	Archaea sp. (clone S
gb_m:AF179926	7.00	105.96	755.59	718	AF179926	Ovis aries SP17 prote	AF179926	Ovis aries SP17 prote	gb_pl1:LUU02467	7.00	104.47	914.90	902	U02467	Lilium longiflorum En
gb_sts:HUMUT7988	7.00	105.95	756.47	719	L30167	Human STS U79988, sequen	L30167	Human STS U79988, sequen	gb_ba2:AF121096	7.00	104.45	917.45	905	AF121096	Uncultured crenarch
gb_pl2:CONS0193W	7.00	105.94	757.35	720	AL11155	Botrytis cinerea strain	AL11155	Botrytis cinerea strain	gb_ba2:AF121093	7.00	104.44	918.30	906	AF121093	Uncultured crenarch
gb_pl2:CONS0198B	7.00	105.94	757.35	720	AL111528	Botrytis cinerea strain	AL111528	Botrytis cinerea strain	gb_ba1:AGLRNNAF	7.00	104.43	919.15	907	M88079	Archaea sp. (clone W
gb_pl2:CONS01977	7.00	105.94	757.35	720	AL111563	Botrytis cinerea strain	AL111563	Botrytis cinerea strain	gb_ba1:AAU11043	7.00	104.42	920.85	909	U11043	Archaeobacterium clone
gb_pl2:CONS019WT	7.00	105.94	757.35	720	AL112186	Botrytis cinerea strain	AL112186	Botrytis cinerea strain	gb_ba2:AF121092	7.00	104.42	920.85	909	AF121092	Uncultured crenarch
gb_pl2:CONS01A01	7.00	105.94	757.35	720	AL112313	Botrytis cinerea strain	AL112313	Botrytis cinerea strain	gb_ba1:AGLRNNA2	7.00	104.41	921.70	910	M88076	Archaea sp. (clone S
gb_pl2:CONS01A0K	7.00	105.94	757.35	720	AL112908	Botrytis cinerea strain	AL112908	Botrytis cinerea strain	gb_ba2:AF121094	7.00	104.41	921.70	910	AF121094	Uncultured crenarch
gb_pl2:CONS01B3U	7.00	105.94	757.35	720	AL113818	Botrytis cinerea strain	AL113818	Botrytis cinerea strain	gb_ba2:AF121098	7.00	104.41	921.70	910	AF121098	Uncultured crenarch
gb_pl2:CONS01B7I	7.00	105.94	757.35	720	AL114670	Botrytis cinerea strain	AL114670	Botrytis cinerea strain	gb_ba2:AF121999	7.00	104.41	922.55	911	AF121999	Uncultured marine a
gb_pl2:CONS01CHT	7.00	105.94	757.35	720	AL115535	Botrytis cinerea strain	AL115535	Botrytis cinerea strain	gb_ba2:AF186421	7.00	104.41	922.55	911	AF186421	Uncultured sponge s
gb_sts:G61833	7.00	105.91	760.88	724	G61833	SHGC-89444 Human Homo sa	G61833	SHGC-89444 Human Homo sa	gb_ba2:AF121993	7.00	104.40	923.39	912	AF121993	Uncultured marine a
gb_pat:CONS01PFW	7.00	105.89	762.64	726	AL142257	Anopheles gambiae STS	AL142257	Anopheles gambiae STS	gb_ba2:AF186422	7.00	104.40	923.39	912	AF186422	Uncultured sponge s
gb_to:100057	7.00	105.86	765.28	729	I00057	Sequence 2 from Patent U	I00057	Sequence 2 from Patent U	gb_ba2:AF186423	7.00	104.40	923.39	912	AF186423	Uncultured sponge s
gb_pat:AR069031	7.00	105.62	788.98	756	AL113493	Botrytis cinerea strain	AL113493	Botrytis cinerea strain	gb_ba1:AF186424	7.00	104.40	923.39	912	AF186424	Uncultured sponge s
gb_pat:EL12900	7.00	105.60	791.60	759	AR069031	Sequence 4 from patent	AR069031	Sequence 4 from patent	gb_ba1:D87350	7.00	104.39	924.24	913	D87350	Unculturable Mariana
gb_pat:194991	7.00	105.60	791.60	759	I94991	Human cDNA encoding HAI-	I94991	Human cDNA encoding HAI-	gb_ba2:AF032943	7.00	104.39	924.24	913	AF032943	Uncultured marine a
gb_pl1:APKVO15	7.00	105.62	799.47	768	X67102	A.phyllitidis myo15 mRNA	X67102	A.phyllitidis myo15 mRNA	gb_ba2:AF122001	7.00	104.39	924.24	913	AF052943	Uncultured marine a
gb_pl2:CONS018W	7.00	105.42	809.93	780	AL10867	Botrytis cinerea strain	AL10867	Botrytis cinerea strain	gb_ba1:ABO15274	7.00	104.38	925.09	914	ABO15274	Unidentified crenar
gb_pl2:CONS019XX	7.00	105.42	809.93	780	AL112021	Botrytis cinerea strain	AL112021	Botrytis cinerea strain	gb_ba1:ABO15276	7.00	104.38	925.09	914	ABO15276	Unidentified crenar
gb_pl2:CONS01CA	7.00	105.42	809.93	780	AL115274	Botrytis cinerea strain	AL115274	Botrytis cinerea strain	gb_ba1:ABO15277	7.00	104.38	925.09	914	ABO15277	Unidentified crenar
gb_pl2:CONS01CHT	7.00	105.42	809.93	780	AL115545	Botrytis cinerea strain	AL115545	Botrytis cinerea strain	gb_ba1:ABO15278	7.00	104.38	925.09	914	ABO15278	Unidentified crenar
gb_ov:AF189767	7.00	105.41	810.80	781	AF189767	Gallus gallus homeodom	AF189767	Gallus gallus homeodom	gb_ba1:D87349	7.00	104.38	925.09	914	D87349	Unculturable Mariana
gb_com:AF181958	7.00	105.39	813.41	784	AF181958	Sus scrofa from Japan	AF181958	Sus scrofa from Japan	gb_ba1:D87349	7.00	104.38	925.09	914	D87349	Unculturable Mariana
gb_com:AF181959	7.00	105.39	813.41	784	AF181959	Sus scrofa from Cook I	AF181959	Sus scrofa from Cook I	gb_ba1:UHU46678	7.00	104.38	925.09	914	U46678	Unidentified hydrothe
gb_com:AF181960	7.00	105.39	813.41	784	AF181960	Sus scrofa from China	AF181960	Sus scrofa from China	gb_ba1:UHU46679	7.00	104.38	925.09	914	U46679	Unidentified

gb_ba2:AF121994	-	7.00	104.38	925.94	915	AF121994	Uncultured marine arch	gb_pl2:AF153765	-	7.00	103.26	1.1e+03	1086	AF153765	Emericella nidulan
gb_ba2:AF121995	-	7.00	104.38	925.94	915	AF121995	Uncultured marine arch	gb_pl1:RTH132C	+	7.00	103.25	1.1e+03	1087	X62459	A.thaliana Hi-2C mrn
gb_ba2:AF052949	-	7.00	104.37	926.79	916	AF052949	Uncultured marine arch	gb_ro:RNU1183A	+	7.00	103.25	1.1e+03	1087	X12934	Rat Ul RNA class II
gb_ba2:AF121991	-	7.00	104.37	926.79	916	AF121991	Uncultured marine arch	gb_v1:IAU204660	+	7.00	103.24	1.1e+03	1089	U20466	Influenza A virus A/
gb_ba2:AF121992	-	7.00	104.37	926.79	916	AF121992	Uncultured marine arch	gb_pr3:ABU59650	+	7.00	103.22	1.1e+03	1092	U59650	Ateles belzebuth MHC
gb_ba2:AF121998	-	7.00	104.37	926.79	916	AF121998	Uncultured marine arch	gb_ba1:AB016843	+	7.00	103.11	1.1e+03	1100	AB016843	Streptomyces thermo
gb_sts:CN501K0A	+	7.00	104.32	932.72	923	AL14723	Anopheles gambiae sts	gb_ov:XU050745	+	7.00	103.17	1.1e+03	1110	U50745	Xenopus laevis homeo
gb_ro:MUSIFNRA	+	7.00	104.26	940.35	932	M28587	Mouse alpha leukocyte in	gb_in1:FPAGBPA	+	7.00	103.02	1.1e+03	1126	M10985	Plasmodium falciparu
gb_pl1:NTA3J177	+	7.00	104.25	941.35	933	AD223177	Nicotiana tabacum mrna	gb_in2:FEU96712	+	7.00	102.98	1.1e+03	1133	U96712	Trypanosoma brucei p
gb_v1:JSU12798	+	7.00	104.24	942.88	933	U12798	Jerry Slough virus BFS44	gb_in2:AF083914	+	7.00	102.96	1.1e+03	1137	AF083914	Arabidopsis thalia
em_pat:EI11385	+	7.00	104.23	943.73	936	EL1385	CDNA encoding mouse neur	gb_pl2:AF083914	+	7.00	102.94	1.1e+03	1140	U05330	Seisseria meningiti
gb_v1:KVU12801	+	7.00	104.22	944.57	937	U12801	Keytong virus C14031-33	gb_pat:108445	+	7.00	102.93	1.1e+03	1141	I08445	Sequencia 1 from Pate
gb_ba1:AB015273	+	7.00	104.19	947.95	941	AF15273	Unidentified crenarcha	gb_ba1:AB015273	+	7.00	102.90	1.1e+03	1147	AF087839	Rattus norvegicus
gb_v1:MVU12802	+	7.00	104.18	947.95	941	U12802	Melao virus TVU19375 S	gb_ro:RNSU3251	+	7.00	102.87	1.1e+03	1152	AF087839	Rattus norvegicus
gb_pl2:AF115527	+	7.00	104.18	949.64	943	AF115527	Arabidopsis thaliana c	gb_pl1:NTA3J178	+	7.00	102.86	1.1e+03	1153	U323178	Nicotiana tabacum
gb_sts:CN501K1K4	+	7.00	104.11	958.08	933	AL145829	Anopheles gambiae sts	gb_ba1:RSU32870	+	7.00	102.86	1.1e+03	1153	U32870	Rhodococcus sp. thcr
gb_pr2:HSU56421	+	7.00	104.10	958.92	934	U56421	Human olfactory receptor	gb_v1:TRA7294	+	7.00	102.85	1.1e+03	1156	AJ007294	Tobacco rattle vir
gb_sts:G31753	+	7.00	104.10	958.92	934	G31753	SWSG3100 Eric D. Green B	gb_v1:ACA005536	+	7.00	102.77	1.1e+03	1169	AJ005536	Adenovirus clinica
gb_v1:AF072392	+	7.00	104.08	961.45	957	AF072392	Influenza A virus (A/C	gb_in2:AF228517	+	7.00	102.77	1.1e+03	1170	AF228517	Drosophila melanog
gb_v1:AF072394	+	7.00	104.08	961.45	957	AF072394	Influenza A virus (A/C	gb_pl1:SPS28P1	+	7.00	102.76	1.1e+03	1171	X59987	S.pombe SSp1 gene fo
gb_v1:AF072395	+	7.00	104.08	961.45	957	AF072395	Influenza A virus (A/C	gb_in1:HYDANEX	+	7.00	102.62	1.2e+03	1197	M83736	Hydra vulgaris annex
gb_v1:AF072396	+	7.00	104.08	961.45	957	AF072396	Influenza A virus (A/C	gb_in1:AFYF81	+	7.00	102.60	1.2e+03	1200	X05578	Japanese oak silkwor
gb_v1:AF072397	+	7.00	104.08	961.45	957	AF072397	Influenza A virus (A/C	gb_ba2:AF005679	+	7.00	102.60	1.2e+03	1200	X05578	Japanese oak silkwor
gb_v1:AF072398	+	7.00	104.08	961.45	957	AF072398	Influenza A virus (A/C	gb_ba1:RCEP50233	+	7.00	102.31	1.2e+03	1256	AF005679	Shewanella putrefa
gb_v1:AF072399	+	7.00	104.08	961.45	957	AF072399	Influenza A virus (A/C	gb_pl1:MCZU1293	+	7.00	102.26	1.2e+03	1264	X55896	E.coli plasmid pSU23
gb_v1:AF072400	+	7.00	104.08	962.29	958	AF072400	Growth hormone [Sparus a	gb_in1:AF044734	+	7.00	102.20	1.2e+03	1277	L35913	Zea mays lipase (LIP
gb_v1:AF072401	+	7.00	104.08	962.29	958	AF072401	Growth hormone [Sparus a	gb_in2:AF044734	+	7.00	102.18	1.2e+03	1281	AF044734	Drosophila melanog
gb_v1:AF072402	+	7.00	104.08	962.29	958	AF072402	Growth hormone [Sparus a	gb_ba1:FPARBA	+	7.00	102.16	1.2e+03	1304	Y10458	P.fluorescens arba g
gb_v1:AF072403	+	7.00	104.08	962.29	958	AF072403	Growth hormone [Sparus a	gb_pl2:LLMEIOTIN	+	7.00	102.00	1.3e+03	1317	L20696	Lilium longiflorum m
gb_v1:AF072404	+	7.00	104.08	962.29	958	AF072404	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	102.00	1.3e+03	1320	X04730	Enterobacter cloacae
gb_v1:AF072405	+	7.00	104.08	962.29	958	AF072405	Growth hormone [Sparus a	gb_ba2:UCU78199	+	7.00	101.98	1.3e+03	1320	U51475	Rhizobium meliloti e
gb_v1:AF072406	+	7.00	104.08	962.29	958	AF072406	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.98	1.3e+03	1320	U51475	Rhizobium meliloti e
gb_v1:AF072407	+	7.00	104.08	962.29	958	AF072407	Growth hormone [Sparus a	gb_in2:AF009328	+	7.00	101.92	1.3e+03	1331	AF009328	Mus musculus orpha
gb_v1:AF072408	+	7.00	104.08	962.29	958	AF072408	Growth hormone [Sparus a	gb_ro:STU83828	+	7.00	101.91	1.3e+03	1334	U83828	Scotinomyces tegula N
gb_v1:AF072409	+	7.00	104.08	962.29	958	AF072409	Growth hormone [Sparus a	gb_ro:MMU48150	+	7.00	101.87	1.3e+03	1343	U48150	Mus musculus male me
gb_v1:AF072410	+	7.00	104.08	962.29	958	AF072410	Growth hormone [Sparus a	gb_ba2:UCU78199	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072411	+	7.00	104.08	962.29	958	AF072411	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072412	+	7.00	104.08	962.29	958	AF072412	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072413	+	7.00	104.08	962.29	958	AF072413	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072414	+	7.00	104.08	962.29	958	AF072414	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072415	+	7.00	104.08	962.29	958	AF072415	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072416	+	7.00	104.08	962.29	958	AF072416	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072417	+	7.00	104.08	962.29	958	AF072417	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072418	+	7.00	104.08	962.29	958	AF072418	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072419	+	7.00	104.08	962.29	958	AF072419	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072420	+	7.00	104.08	962.29	958	AF072420	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072421	+	7.00	104.08	962.29	958	AF072421	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072422	+	7.00	104.08	962.29	958	AF072422	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072423	+	7.00	104.08	962.29	958	AF072423	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072424	+	7.00	104.08	962.29	958	AF072424	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072425	+	7.00	104.08	962.29	958	AF072425	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072426	+	7.00	104.08	962.29	958	AF072426	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072427	+	7.00	104.08	962.29	958	AF072427	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072428	+	7.00	104.08	962.29	958	AF072428	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072429	+	7.00	104.08	962.29	958	AF072429	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072430	+	7.00	104.08	962.29	958	AF072430	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072431	+	7.00	104.08	962.29	958	AF072431	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072432	+	7.00	104.08	962.29	958	AF072432	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072433	+	7.00	104.08	962.29	958	AF072433	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072434	+	7.00	104.08	962.29	958	AF072434	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072435	+	7.00	104.08	962.29	958	AF072435	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072436	+	7.00	104.08	962.29	958	AF072436	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072437	+	7.00	104.08	962.29	958	AF072437	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072438	+	7.00	104.08	962.29	958	AF072438	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072439	+	7.00	104.08	962.29	958	AF072439	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072440	+	7.00	104.08	962.29	958	AF072440	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072441	+	7.00	104.08	962.29	958	AF072441	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072442	+	7.00	104.08	962.29	958	AF072442	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072443	+	7.00	104.08	962.29	958	AF072443	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072444	+	7.00	104.08	962.29	958	AF072444	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc
gb_v1:AF072445	+	7.00	104.08	962.29	958	AF072445	Growth hormone [Sparus a	gb_ba1:RCEP50233	+	7.00	101.85	1.3e+03	1346	U78199	Unidentified crenarc

gb_prl:AF135016	1.4e+03	7.00	101.22	1482	1	AF135016 Homo sapiens protein	gb_ov:AF165889	7.00	99.37	1.8e+03	1967
gb_prl:HUMNF1A	1.4e+03	7.00	101.22	1483	1	M61213 Human neurofibromatosis 1	gb_in2:AF067126	7.00	99.25	1.8e+03	2005
gb_prl:AB006534	1.4e+03	7.00	101.09	1512	1	AB006534 Homo sapiens mRNA for	gb_prl:GHMALS	7.00	99.23	1.8e+03	2010
gb_bal:GLUCYTTOCH	1.4e+03	7.00	101.06	1519	1	M58760 G suboxydants cytochrome	gb_inl:DROCACTUSX	7.00	99.23	1.8e+03	2011
gb_pll1:2MDJ36437	1.4e+03	7.00	101.01	1529	1	U34337 Zea mays H+pyrophosphatase	gb_bal:AB001720	7.00	99.21	1.8e+03	2015
gb_ov:RNNARGCON	1.4e+03	7.00	101.02	1530	1	AF000349 Rattus norvegicus N-ras	gb_ov:MM1FN55	7.00	99.20	1.8e+03	2021
gb_prl2:VCPHEI	1.4e+03	7.00	100.98	1539	1	X59401 Volvox carteri mRNA for	gb_bal:SSM04198	7.00	99.13	1.8e+03	2040
gb_inl:TEMAPK	1.4e+03	7.00	100.96	1543	1	Z53431 T.brucei gene for MAP K	gb_prl:HSR801199	7.00	99.12	1.8e+03	2043
gb_prl:HSU078095	1.4e+03	7.00	100.90	1544	1	U78095 Homo sapiens placental	gb_inl:DROCACTUSB	7.00	99.08	1.8e+03	2057
gb_vl:AB009268	1.4e+03	7.00	100.96	1557	1	AB009268 Human herpesvirus 1 U	gb_bal:AFATPDA	7.00	99.08	1.8e+03	2058
gb_bal:SZU25852	1.4e+03	7.00	100.89	1563	1	U25852 Streptococcus zooepidemicus	gb_prl:AL12337	7.00	99.08	1.8e+03	2058
gb_pll1:AB012103	1.5e+03	7.00	100.88	1566	1	AB012103 Triticum aestivum mRNA	gb_prl:AL136756	7.00	99.08	1.8e+03	2058
gb_prl:AB0272405	1.5e+03	7.00	100.87	1564	1	AF027205 Homo sapiens Kunitz-b	gb_prl:AF011377	7.00	99.07	1.8e+03	2059
gb_bal:DS85038	1.5e+03	7.00	100.87	1565	1	D85038 Caldococcus noboribetanus	gb_prl:HSAL132261	7.00	99.06	1.8e+03	2063
gb_prl:HSU083659	1.5e+03	7.00	100.87	1565	1	U83659 Human multidrug resistanc	gb_ov:MUSNYI8B	7.00	99.05	1.8e+03	2065
gb_prl1:CHCE1165	1.5e+03	7.00	100.79	1583	1	U12742 Chlorella ellipsoidea c	gb_inl:AE1663314	7.00	99.05	1.8e+03	2066
gb_ov:XLNOVA	1.5e+03	7.00	100.75	1593	1	Z85983 Xenopus laevis mRNA for	gb_prl:EL15324	7.00	98.96	1.9e+03	2096
gb_prl:AF054986	1.5e+03	7.00	100.75	1594	1	AF054986 Homo sapiens clone 23	em_pat:EL11883	7.00	98.96	1.9e+03	2096
gb_ov:AF081350	1.5e+03	7.00	100.73	1597	1	AF081350 Xenopus laevis homeod	em_pat:EL11883	7.00	98.94	1.9e+03	2102
gb_ov:XLACHRB	1.5e+03	7.00	100.73	1599	1	U04618 Xenopus laevis muscle n	gb_inl:CHXPB03	7.00	98.94	1.9e+03	2106
gb_ov:MMPCAD	1.5e+03	7.00	100.72	1600	1	X68057 M.musculus gene for P-c	gb_prl:AHHL2FLK	7.00	98.93	1.9e+03	2106
gb_bal:NGAZURIN	1.5e+03	7.00	100.68	1610	1	X91654 P.putida nlpD (partial)	gb_in2:DROLACH	7.00	98.91	1.9e+03	2110
gb_prl:HSM801262	1.5e+03	7.00	100.63	1624	1	AL122105 Homo sapiens mRNA; c	gb_prl:NCU78026	7.00	98.86	1.9e+03	2127
gb_bal:PAE15166	1.5e+03	7.00	100.61	1628	1	U15166 Pseudomonas aeruginosa	gb_in2:AF107698	7.00	98.83	1.9e+03	2136
gb_bal:PAE18494	1.5e+03	7.00	100.61	1628	1	Y18494 Pseudomonas aeruginosa	gb_prl:AF129264	7.00	98.82	1.9e+03	2139
gb_vl:AF071775	1.5e+03	7.00	100.61	1629	1	AF071775 Influenza A virus str	gb_bal:ECHELYL	7.00	98.81	1.9e+03	2144
gb_vl:AF071776	1.5e+03	7.00	100.61	1629	1	AF071776 Influenza A virus str	gb_prl:AF113249	7.00	98.75	1.9e+03	2163
gb_bal:PPRPO5	1.5e+03	7.00	100.59	1634	1	X3240 Human enigma gene, comp	gb_ov:RNSVFG	7.00	98.75	1.9e+03	2165
gb_bal:BSPO10057	1.5e+03	7.00	100.47	1662	1	AF010057 Burkholderia sp. JH1	gb_bal:PPU010261	7.00	98.74	1.9e+03	2168
gb_bal:RHMEXXOY	1.5e+03	7.00	100.46	1665	1	M61751 R.meliloti ExoX protein	gb_prl:HSU93822	7.00	98.71	1.9e+03	2176
gb_ov:XLFLRFO5	1.5e+03	7.00	100.43	1674	1	M6427 Murine lactoferrin gene	gb_prl:SOYCPGASAT	7.00	98.70	1.9e+03	2181
gb_ov:BTGABR3	1.6e+03	7.00	100.31	1704	1	X12362 Bovine mRNA for GABA-A	gb_bal:BCACPSAT	7.00	98.69	1.9e+03	2185
gb_pll1:YSCRNABIND	1.6e+03	7.00	100.29	1709	1	L01797 Saccharomyces cerevisiae	gb_prl:HUMRSU2	7.00	98.66	1.9e+03	2194
gb_bal:TXLYLA	1.6e+03	7.00	100.28	1711	1	Z81013 T.alba xyla gene. 10/19	gb_ov:RNSNATP2	7.00	98.65	1.9e+03	2197
gb_prl:HUMENIGMA	1.6e+03	7.00	100.23	1725	1	X3240 Human enigma gene, comp	gb_ov:RNSNATP2	7.00	98.65	1.9e+03	2197
gb_prl:ATASKRNA	1.6e+03	7.00	100.18	1738	1	X94939 A.thaliana mRNA for	gb_ov:RNSNATP2	7.00	98.54	2.0e+03	2235
gb_prl:HUMES	1.6e+03	7.00	100.16	1745	1	D64016 Human gene for vascular	gb_prl:USMORN50XB	7.00	98.52	2.0e+03	2240
gb_prl2:AF096247	1.6e+03	7.00	100.13	1751	1	AF096247 Lycopersicon esculent	gb_prl:AF154001	7.00	98.52	2.0e+03	2242
gb_ov:AF096269	1.6e+03	7.00	100.13	1752	1	AF096269 Rattus norvegicus EH	gb_bal:FUMH1185	7.00	98.49	2.0e+03	2250
gb_ov:CSPTPIN	1.6e+03	7.00	100.07	1768	1	X14898 Hamster p7 preinsertion	gb_vl:HSPIRA4	7.00	98.49	2.0e+03	2250
gb_vl:MLGEP55	1.6e+03	7.00	100.05	1774	1	J02193 friend spleen focus-for	gb_prl:AF095794	7.00	98.49	2.0e+03	2252
gb_bal2:AF047381	1.6e+03	7.00	100.03	1780	1	AF047381 Pseudomonas aeruginos	gb_prl:PTUBB4GN	7.00	98.48	2.0e+03	2254
gb_inl:AK118RNA	1.6e+03	7.00	99.97	1796	1	Z12671 A.kurogane 18S ribosoma	gb_prl:YSCPBUX1X	7.00	98.48	2.0e+03	2256
gb_inl:TS060231	1.6e+03	7.00	99.96	1798	1	U00231 Trichinella spiralis 18	gb_ov:AF029758	7.00	98.46	2.0e+03	2261
gb_inl:AB005566	1.6e+03	7.00	99.96	1799	1	AB005566 Bombyx mori mRNA for	gb_bal:D873233	7.00	98.46	2.0e+03	2263
gb_bal2:A26178	1.6e+03	7.00	99.94	1803	1	A26178 T. gondii Tg34 cDNA c	gb_prl:SGC13BG	7.00	98.33	2.0e+03	2306
gb_inl:SS4994	1.6e+03	7.00	99.94	1804	1	S54994 54-kda antigen [Toxoplas	gb_prl:LSGAB9	7.00	98.27	2.0e+03	2327
gb_inl:DMSCPX	1.6e+03	7.00	99.94	1806	1	X97685 D.melanogaster mRNA for	gb_prl:DS22935	7.00	98.23	2.0e+03	2343
gb_ov:AB006452	1.6e+03	7.00	99.93	1806	1	AB006452 Rattus norvegicus mRN	gb_prl:HSPIBDA	7.00	98.21	2.0e+03	2349
gb_bal:XXANAMY	1.6e+03	7.00	99.91	1813	1	M85252 Xanthomonas campestris	gb_prl:RNUTP1AC	7.00	98.16	2.1e+03	2366
gb_ov:KELDPFPA	1.6e+03	7.00	99.88	1818	1	M23238 X.laevis platelet-deriv	gb_inl:AAU06475	7.00	98.15	2.1e+03	2373
gb_pll1:NTPTVP17	1.7e+03	7.00	99.82	1838	1	X83728 N.tabacum mRNA for inor	gb_bal:FVBEND08EP	7.00	98.05	2.1e+03	2408
gb_pll1:ACWARS017	1.7e+03	7.00	99.79	1840	1	X93081 A.immersus putative gen	gb_vl:MCU86887	7.00	98.04	2.1e+03	2411
gb_pll1:SCYB0744	1.7e+03	7.00	99.77	1854	1	Z35943 S.cerevisiae chromosome	gb_bal:PDEENFAB	7.00	97.95	2.1e+03	2440
gb_pll1:MDRBSACA	1.7e+03	7.00	99.76	1853	1	Z21794 M.domestica ribulose-1,	gb_prl:DS22935	7.00	97.95	2.1e+03	2445
gb_prl:HSR801376	1.7e+03	7.00	99.69	1874	1	X133099 Homo sapiens mRNA; c	gb_ov:RNNWTP2	7.00	98.21	2.1e+03	2448
gb_bal:D86856	1.7e+03	7.00	99.69	1875	1	D68566 Bacillus subtilis soda	gb_ov:OMRABEXR	7.00	97.94	2.1e+03	2484
gb_bal:CJ5224793	1.7e+03	7.00	99.68	1876	1	AF224793 Campylobacter jejuni	gb_prl:HSACSTR	7.00	97.92	2.1e+03	2457
gb_ov:MAU26454	1.7e+03	7.00	99.64	1887	1	U26454 Mesocricetus auratus ec	gb_ov:MUSHR5	7.00	97.92	2.1e+03	2458
gb_pll1:SCU12311	1.7e+03	7.00	99.77	1892	1	U12311 Saccharomyces cerevisia	gb_ov:MUSINIBRA	7.00	97.91	2.1e+03	2459
gb_prl:HS4223957	1.7e+03	7.00	99.62	1893	1	AF223357 Homo sapiens mRNA for	gb_prl:TUJ35436	7.00	97.91	2.1e+03	2460
gb_inl:DROCACTUSA	1.7e+03	7.00	99.61	1896	1	L03367 Drosophila melanogaster	gb_prl:DB9927	7.00	97.89	2.1e+03	2465
gb_prl:HSR801309	1.7e+03	7.00	99.60	1900	1	AB003109 Humicola grisea var.	gb_prl:ARHTWARS1	7.00	97.89	2.1e+03	2469
gb_ov:MUSFCRAC	1.7e+03	7.00	99.56	1912	1	D13547 Mus musculus DNA, T ear	gb_pat:AR007280	7.00	97.87	2.1e+03	2476
gb_pat:I120298	1.7e+03	7.00	99.54	1921	1	E12528 DNA encoding an enzyme	gb_bal:AR062699	7.00	97.84	2.1e+03	2485
gb_prl:136611	1.7e+03	7.00	99.53	1921	1	I20390 Sequence 28 from patent	gb_bal:RMRP5A	7.00	97.82	2.1e+03	2495
gb_ov:KELDPGFA	1.7e+03	7.00	99.47	1939	1	M23237 X.laevis platelet-deriv	gb_ov:CHRCAPZA	7.00	97.80	2.2e+03	2500
gb_pll1:SCYB081W	1.7e+03	7.00	99.45	1945	1	Z35842 S.cerevisiae chromosome	gb_prl:SCYOB190W	7.00	97.80	2.2e+03	2502
gb_inl:ECRRN18	1.8e+03	7.00	99.41	1957	1	X13457 Bird spider 18S ribosom	gb_inl:AHNDTHY	7.00	97.80	2.2e+03	2503
gb_bal:AF102982	1.8e+03	7.00	99.40	1958	1	AB010282 Mus musculus mRNA for	gb_prl:VKN01267	7.00	97.75	2.2e+03	2522
gb_prl:AF124143	1.8e+03	7.00	99.40	1958	1	AF124143 Homo sapiens cyclin A	gb_prl:SCALG9GN	7.00	97.74	2.2e+03	2522
gb_ov:E07956	1.8e+03	7.00	99.39	1962	1	EF07956 cDNA encoding transglut	gb_prl:SCYNL219C	7.00	97.74	2.2e+03	2522
AF165889 Gallus gallus inhi											
AF067126 Bemisia argentifol											
X52305 Cotton mRNA for mala											
L04964 Drosophila melanogae											
AB001720 Borrelia turicatae											
XU1971 Mouse gene for inter											
U04088 Streptomyces sp. lac											
AL117661 Homo sapiens mRNA;											
L03368 Drosophila melanogae											
M16730 A.eutrophus 2,4-dich											
A12337 ttda gene cloned in											
I36756 Sequence 1 from pate											
AF011377 Toxoplasma gondii											
AL132261 Homo sapiens mRNA											
D83819 Mouse mRNA for neuro											
U65314 Aedes aegypti allato											
E15324 Human mRNA for L-as											
E11883 cDNA encoding human											
XU19795 C. erythrocephala gen											
X2460 A.thaliana gene for											
L13255 Fruitfly lachesis mR											
AF107698 Aedes aegypti clone											
AF129264 Homo sapiens clon											
XU1047 E.coli gene for ente											
AF113249 Homo sapiens multi											
XU1958 Rat seminal vesicle											
AU010261 Pseudomonas putida											
U95822 Human putative trans											
L22964 Glycine soja chlorop											
Z46432 Bacillus circulans s											
L12335 Human RSU-1/RSP-1 mR											
XU175305 R.norvegicus (PVG-R											
XU175307 R.norvegicus (PVG-R											
Z36906 T.gondii gene for rh											
M95233 Corn smut orthinine-											
AF154001 Homo sapiens MRP3											
U81185 Fusobacterium mortif											
M76726 Equine herpesvirus t											
AF095794 Schizosaccharomyce								</			

gb_inl:DMNOTB	+	7.00	97.74	2.2e+03	2526	X95248 D.melanogaster l(2)not	gb_bal:BPE7362	7.00	96.72	2.5e+03	2949	AJ007362 Bordetella pertussis
gb_inl:DMNOT4M	+	7.00	97.74	2.2e+03	2526	X95252 D.melanogaster l(2)not	gb_bal:MTBRLC2	7.00	96.72	2.5e+03	2951	AF189751 Mycobacterium tube
gb_bal:RSMOB	+	7.00	97.74	2.2e+03	2527	Y05660 R.sphaeroides mob gene	gb_bal:FBVAPPE	7.00	96.72	2.5e+03	2952	D10980 Flavobacterium menin
gb_pll:ADG09077	+	7.00	97.72	2.2e+03	2531	AB009077 Vigna radiata mRNA fc	gb_pll:ADG07369	7.00	96.71	2.5e+03	2956	AF017369 Mus musculus facio
gb_inl:ZDGRAND16	+	7.00	97.72	2.2e+03	2533	X97605 Z.diploperennis DNA for	gb_pll:OSGT2	7.00	96.71	2.5e+03	2957	X68261 O.sativa gt-2 gene
gb_pl1:NTITVTP9	+	7.00	97.67	2.2e+03	2551	X83730 N.tabacum mRNA for inon	gb_inl:DMU89264	7.00	96.69	2.5e+03	2963	AJ007361 Bordetella pertussis
gb_pl2:SCU23759	+	7.00	97.67	2.2e+03	2551	X23759 Saccharomyces cerevisia	gb_inl:DMU89264	7.00	96.69	2.5e+03	2963	U89264 Drosophila melanogas
gb_pr3:AF058289	+	7.00	97.67	2.2e+03	2551	AF058289 Homo sapiens type II	gb_bal:FMPE1PEP1	7.00	96.69	2.5e+03	2964	X63674 F.meningosepticum fl
gb_pr3:HSM080173	+	7.00	97.65	2.2e+03	2558	AF117636 Homo sapiens mRNA; CL	gb_ro:MUSNRAMPB	7.00	96.67	2.5e+03	2972	L33415 Mouse integral membr
gb_pr5:AF176097	+	7.00	97.65	2.2e+03	2561	AF176097 Homo sapiens DNA; po	gb_ro:MUSNRAMPB	7.00	96.67	2.5e+03	2972	L33415 Mouse integral membr
gb_inl:DMRHAH15	+	7.00	97.62	2.2e+03	2563	X98766 D.melanogaster mRNA for	gb_pll:SCYNL016W	7.00	96.61	2.5e+03	2998	X71292 S.cerevisiae chromos
gb_pr5:AF113676	+	7.00	97.62	2.2e+03	2571	AF113676 Homo sapiens clone FL	gb_bal:BBPRNG	7.00	96.61	2.5e+03	3000	X54815 B.bronchiseptica pr
gb_ov:CHK130ES	+	7.00	97.58	2.2e+03	2582	L13973 Gallus gallus EDT-solub	gb_bal:BBPRNG	7.00	96.61	2.5e+03	3000	X54815 B.bronchiseptica pr
gb_pr3:HSAL131891	+	7.00	97.58	2.2e+03	2589	AJ131891 Homo sapiens mRNA for	gb_bal:AF004221	7.00	96.61	2.5e+03	3000	AF04560 B.pertussis omp69A g
gb_bal:XCPRUGB	+	7.00	97.49	2.2e+03	2623	Z37113 X.campestris fruB gene	gb_bal:AF004221	7.00	96.61	2.5e+03	3000	AF04560 B.pertussis omp69A g
gb_bal:AF47683	+	7.00	97.46	2.2e+03	2636	A47683 Sequence 1 from Patent	gb_pat:AL9178	7.00	96.61	2.5e+03	3000	AF04560 B.pertussis omp69A g
gb_pat:121366	+	7.00	97.46	2.2e+03	2636	I21366 Sequence 1 from Patent	gb_pat:AL9180	7.00	96.61	2.5e+03	3000	AF04560 B.pertussis omp69A g
gb_ro:RATPRCCG1	+	7.00	97.45	2.3e+03	2639	K03243 Rat phosphoenolpyruvate	gb_pat:A26124	7.00	96.61	2.5e+03	3000	AF04560 B.pertussis omp69A g
gb_pl1:BLV1NOPP	+	7.00	97.43	2.3e+03	2649	D13472 Hordeum vulgare mRNA fc	gb_pat:A26124	7.00	96.61	2.5e+03	3000	AF04560 B.pertussis omp69A g
gb_in2:SUSTEKA1	+	7.00	97.41	2.3e+03	2656	M97188 Strongylocentrotus purp	gb_in2:CAU2531465	7.00	96.61	2.5e+03	3002	AF099741 Penaeus monodon pr
gb_pl2:YSCPRUBIP	+	7.00	97.35	2.3e+03	2675	L20767 Saccharomyces cerevisia	gb_in2:CAU2531465	7.00	96.61	2.5e+03	3002	AF099741 Penaeus monodon pr
gb_in2:AF145667	+	7.00	97.32	2.3e+03	2691	AF145667 Drosophila melanogast	gb_ro:AF042066	7.00	96.55	2.5e+03	3027	AF042066 Mus musculus domes
gb_bal:FPFPCOAH1	+	7.00	97.30	2.3e+03	2700	Y13067 Pseudomonas fluorescens	gb_pr5:AF119256	7.00	96.53	2.5e+03	3038	AF119256 Homo sapiens PAT2
gb_pl1:D86306	+	7.00	97.29	2.3e+03	2704	D86306 Cucurbita moschata mRNA	gb_pr4:HSCTSZ1	7.00	96.51	2.5e+03	3046	AF136274 Homo sapiens cathe
gb_pl1:D45384	+	7.00	97.28	2.3e+03	2710	D45384 Oryza sativa mRNA for v	gb_ro:AF109674	7.00	96.50	2.5e+03	3052	AF109674 Rattus norvegicus
gb_pl1:NTSAMIPP	+	7.00	97.27	2.3e+03	2713	X77915 N.tabacum mRNA for inon	gb_in2:HMU36781	7.00	96.50	2.5e+03	3052	AF109674 Rattus norvegicus
gb_bal:BPE011015	+	7.00	97.26	2.3e+03	2718	AJ011015 Bordetella pertussis	gb_vi:MLXENVXA	7.00	96.48	2.5e+03	3060	K02730 Murine leukemia viru
gb_bal:BPE011016	+	7.00	97.26	2.3e+03	2718	AJ011016 Bordetella pertussis	gb_in2:AF188718	7.00	96.46	2.6e+03	3069	AF188718 Drosophila melanog
gb_bal:BPE133245	+	7.00	97.26	2.3e+03	2718	AJ133245 Bordetella pertussis	gb_ro:MMU23235	7.00	96.44	2.6e+03	3080	U23235 Mus musculus facio
gb_vi:ADRV7RHND	+	7.00	97.26	2.3e+03	2718	M38648 Mastadenovirus 7 HindII	gb_bal:ECN1DLFR	7.00	96.40	2.6e+03	3099	Z50802 Escherichia coli dfr
gb_pr3:HSU33920	+	7.00	97.26	2.3e+03	2719	X3920 Human clone lambda 5 se	gb_pl1:AMFTLIG	7.00	96.40	2.6e+03	3101	X57296 A.majus filli gene fo
gb_bal:BPE132095	+	7.00	97.24	2.3e+03	2724	AJ132095 Bordetella pertussis	gb_ro:AF021806	7.00	96.37	2.6e+03	3115	AF021806 Rattus norvegicus
gb_bal:BPE133784	+	7.00	97.22	2.3e+03	2733	AJ133784 Bordetella pertussis	gb_bal:BPPRAN	7.00	96.36	2.6e+03	3116	X54547 Bordetella parapertu
gb_in2:AF149795	+	7.00	97.22	2.3e+03	2735	AF149795 Drosophila melanogast	gb_in2:AF188718	7.00	96.36	2.6e+03	3117	AF188718 Drosophila melanog
gb_pr4:HSGLVRK61	+	7.00	97.19	2.3e+03	2747	AF102059 Homo sapiens gibbon a	gb_in2:AF188718	7.00	96.32	2.6e+03	3136	AF051695 Trypanosoma cruzi
gb_pl1:HSGLVRK61	+	7.00	97.16	2.3e+03	2757	AB016896 Schizosaccharomyces f	gb_in2:AF051695	7.00	96.30	2.6e+03	3146	AF051695 Trypanosoma cruzi
gb_ro:RUH07863	+	7.00	97.12	2.4e+03	2775	AB016896 Schizosaccharomyces f	gb_in2:AF051695	7.00	96.29	2.6e+03	3154	AF051695 Trypanosoma cruzi
gb_pr1:AB018551	+	7.00	97.10	2.4e+03	2785	AB018551 Homo sapiens ATPBL m	gb_pr1:AB008855	7.00	96.28	2.6e+03	3156	AB008855 Arabidopsis thalia
gb_in2:AF035546	+	7.00	97.08	2.4e+03	2795	AF035546 Drosophila melanogast	gb_pl1:AB008855	7.00	96.28	2.6e+03	3157	AB008855 Arabidopsis thalia
gb_bal:ACN1FA	+	7.00	97.06	2.4e+03	2801	L32792 Beta vulgaris clone P1	gb_bal:AB008855	7.00	96.28	2.6e+03	3157	AB008855 Arabidopsis thalia
gb_pl1:ATHAVP3	+	7.00	97.04	2.4e+03	2809	X08014 Acorhizobium caulinodan	gb_bal:NGRFBABD	7.00	96.21	2.6e+03	3190	Z32742 N.gonorrhoeae (MS11)
gb_bal:BPE011091	+	7.00	97.03	2.4e+03	2813	AJ011091 Bordetella pertussis	gb_pl1:HSU72761	7.00	96.19	2.7e+03	3200	M89908 Saccharomyces cerevi
gb_bal:BPE011093	+	7.00	97.03	2.4e+03	2814	AJ011093 Bordetella pertussis	gb_pl1:HSU72761	7.00	96.17	2.7e+03	3211	L11794 Sequence 1 from Pat
gb_bal:D38600	+	7.00	97.03	2.4e+03	2816	D38600 Alteromonas sp. DNA for	gb_pr3:HSAC00980	7.00	96.15	2.7e+03	3220	L20859 Human leukemia virus
gb_bal:PE011092	+	7.00	97.02	2.4e+03	2820	AJ01092 Bordetella pertussis	gb_bal:SERATTXIS	7.00	96.09	2.7e+03	3249	AC00980 Homo sapiens (subc
gb_ro:MMCD18	+	7.00	97.00	2.4e+03	2828	X14951 M.musculus mRNA for CDI	gb_pr2:HSRFA60	7.00	96.04	2.7e+03	3276	L11597 Saccharopolyspora s
gb_pr4:AF03480	+	7.00	96.99	2.4e+03	2830	E03480 cDNA encoding rice lip	gb_pr3:HSRFA60	7.00	96.04	2.7e+03	3276	L11597 Saccharopolyspora s
gb_pl1:OSLRNA	+	7.00	96.99	2.4e+03	2830	X64396 O.sativa mRNA for lip	gb_pr3:HSRFA60	7.00	96.04	2.7e+03	3276	L11597 Saccharopolyspora s
gb_ro:AF022136	+	7.00	96.99	2.4e+03	2833	AF022136 Rattus norvegicus con	gb_pr2:HSU18288	7.00	95.87	2.8e+03	3346	M76981 Glycine max vspA gen
gb_pl1:YSCWH12	+	7.00	96.98	2.4e+03	2837	M21089 Yeast (S.cerevisiae) g	gb_pl1:SCVORLYWAX	7.00	95.87	2.8e+03	3360	Z74974 S.cerevisiae chromos
gb_ov:AB008372	+	7.00	96.94	2.4e+03	2854	AB008372 Oncorhynchus mykiss m	gb_bal:BNBSLAG	7.00	95.86	2.8e+03	3363	U18288 Human clone CIITA-10
gb_ro:HAMNFA	+	7.00	96.94	2.4e+03	2854	J04122 Hamster nuclear factor	gb_bal:BNBSLAG	7.00	95.82	2.8e+03	3388	L82000 Rhodobacter capsulat
gb_pr3:HSMD02238	+	7.00	96.93	2.4e+03	2856	AL137722 Homo sapiens mRNA; CD	gb_pr3:HSMD02238	7.00	95.82	2.8e+03	3389	U83704 Mesocricetus auratus
gb_pr2:HUMTHC	+	7.00	96.91	2.4e+03	2866	M23597 Homo sapiens tyrosine h	gb_pat:167678	7.00	95.81	2.8e+03	3393	I67678 Sequence 1 from pate
gb_pl1:NTITVTP31	+	7.00	96.91	2.4e+03	2867	X83729 N.tabacum mRNA for inon	gb_pl1:CAU18993	7.00	95.77	2.8e+03	3412	U18983 Candida albicans adh
gb_bal:BEUPFRO	+	7.00	96.91	2.4e+03	2868	L32791 Beta vulgaris clone P2	gb_bal:ECOCARD9	7.00	95.76	2.8e+03	3419	U12598 Escherichia coli K-1
gb_bal:A92133	+	7.00	96.90	2.4e+03	2871	A92133 Sequence 2 from Patent	gb_pr4:AF088916	7.00	95.73	2.8e+03	3431	AF088916 Homo sapiens emil
gb_bal:BPA006160	+	7.00	96.84	2.4e+03	2898	AJ006160 Bordetella pertussis	gb_bal:MTBRLM2	7.00	95.73	2.8e+03	3432	AF189826 Mycobacterium bovi
gb_ov:SPSUD	+	7.00	96.83	2.4e+03	2901	U27454 Lycopersicon esculentum	gb_ro:MAU33705	7.00	95.72	2.8e+03	3440	U83705 Mesocricetus auratus
gb_pr4:AF129263	+	7.00	96.82	2.4e+03	2906	AF129263 Homo sapiens clone BF	gb_bal:STU30491	7.00	95.72	2.8e+03	3441	U30491 Salmonella typhimuri
gb_bal:AF107095	+	7.00	96.81	2.4e+03	2910	AF107095 Rhodobacter sphaeroid	gb_pr3:HSMD00438	7.00	95.66	2.8e+03	3469	AF050138 Homo sapiens mRNA
gb_pl1:CAU44747	+	7.00	96.80	2.4e+03	2913	U44747 Candida albicans adhesi	gb_inl:DMNOT56	7.00	95.60	2.9e+03	3500	X77820 D.melanogaster mRNA
gb_ro:RNRSTAT3PT	+	7.00	96.78	2.5e+03	2924	X18180 R.norvegicus mRNA for S	gb_pl1:SCUR154W	7.00	95.59	2.9e+03	3504	Z49654 S.cerevisiae chromos
gb_bal:MLRSCA	+	7.00	96.77	2.5e+03	2927	M12897 Plasmodium falciparum q	gb_om:RABEPD	7.00	95.59	2.9e+03	3507	D13310 Rabbit mRNA for endo
gb_vi:EMLVGP70	+	7.00	96.75	2.5e+03	2936	X53824 Endogenous murine leuk	gb_pr1:AB003592	7.00	95.55	2.9e+03	3530	AB003592 Homo sapiens mRNA
gb_bal:BPA006152	+	7.00	96.75	2.5e+03	2939	AJ006152 Bordetella pertussis	gb_ro:MMU31908	7.00	95.53	2.9e+03	3538	U31908 Mus musculus potassi
gb_bal:STYRIGH	+	7.00	96.73	2.5e+03	2947	M24466 S.typhimurium flagellar	gb_bal:PCBDBABC	7.00	95.52	2.9e+03	3548	X79076 P.cepacia (2CBS) cbd
gb_bal:BPA006156	+	7.00	96.72	2.5e+03	2949	AJ006156 Bordetella pertussis	gb_bal:PCBDBABC	7.00	95.51	2.9e+03	3552	AF091240 Acinetobacter sp.
gb_bal:BPA006158	+	7.00	96.72	2.5e+03	2949	AJ006158 Bordetella pertussis	gb_pr2:HSU38276	7.00	95.51	2.9e+03	3552	U38276 Human semaphorin III
							gb_bal:STYRIGH	7.00	95.48	2.9e+03	3566	Z74051 S.cerevisiae chromos
							gb_bal:AV271819	7.00	95.39	2.9e+03	3617	AD271819 Anabaena variabil
							gb_om:BTACHRA1	7.00	95.39	2.9e+03	3618	X02509 B.Taurus mRNA for ac

gb_bal:STU25631 + 7.00 95.38 2.9e+03 3622 ! U25631 Salmonella typhimurium
 gb_sts:G06760 - 7.00 95.31 3.0e+03 3654 ! G06760 human STS WI-7916. 10/1
 gb_bal:PPX12654 - 7.00 95.29 3.0e+03 3670 ! Y12654 P.putida oxo gene. 5/1
 gb_ov:XL081291 + 7.00 95.26 3.0e+03 3689 ! U81291 Xenopus laevis oviducti
 gb_pl1:SCYKL050C + 7.00 95.24 3.0e+03 3701 ! Z8050 S.cerevisiae chromosome
 gb_ba2:AF061029 + 7.00 95.22 3.0e+03 3714 ! AF061029 Pseudomonas syringae
 gb_ba2:HIU32470 + 7.00 95.17 3.0e+03 3740 ! U32470 Haemophilus influenzae
 gb_pr3:HSU88666 - 7.00 95.16 3.0e+03 3745 ! U88666 Homo sapiens serine kin
 gb_pl1:SCU15653 + 7.00 95.15 3.0e+03 3752 ! U15653 Saccharomyces cerevisiae
 gb_ba2:RSU26421 - 7.00 95.15 3.0e+03 3753 ! U26421 Rhodococcus erythropolis
 gb_htg4:AC014184 - 7.00 95.13 3.0e+03 3763 ! AC014184 Drosophila melanogaster

seq_name: gb_pl1:YSCAMY

seq_documentation_block:
 LOCUS YSCAMY 2272 bp DNA PLN 05-JUL-1994
 DEFINITION Saccharomyces cerevisiae aminopeptidase Y gene, complete cds.
 ACCESSION L31635
 VERSION L31635.1 GI:469463
 KEYWORDS aminopeptidase Y.
 SOURCE Saccharomyces cerevisiae (strain D273-10B) DNA.
 ORGANISM Saccharomyces cerevisiae
 Eukaryota; Fungi; Ascomycota; Saccharomycetales;
 Saccharomycetaceae; Saccharomycetes.
 REFERENCE 1 (bases 1 to 2272)
 Nishizawa, M., Yasuhara, T., Nakai, T., Fujiki, Y. and Ohashi, A.
 AUTHORS Molecular cloning of the aminopeptidase Y gene of Saccharomyces
 TITLE cerevisiae. Sequence analysis and gene disruption of a new
 aminopeptidase
 JOURNAL J. Biol. Chem. 269 (18), 13651-13655 (1994)
 MEDLINE 94230479
 FEATURES
 Location/Qualifiers
 1..2272
 /organism="Saccharomyces cerevisiae"
 /strain="D273-10B"
 /db_xref="taxon:4932"
 /map="chromosome II"
 121..126
 TATA_signal
 220..282
 sig_peptide
 220..1833
 CDS
 /SC_number="3.4.11"
 /note="amino acid feature: potential N-glycosylation site
 1, bp 472..474; amino acid feature: potential
 N-glycosylation site 2, bp 505..507; amino acid feature:
 potential N-glycosylation site 3, bp 562
 564; amino
 acid feature: potential N-glycosylation site 4, bp 667..
 669; amino acid feature: potential N-glycosylation site 5,
 bp 703..705; amino acid feature: potential
 N-glycosylation site 6, bp 1330..1332; amino acid
 feature: potential N-glycosylation site 7, bp 1498..
 1500; amino acid feature: potential N-glycosylation site
 8, bp 1657..1659; prepropeptide: bp 220..387"
 /codon_start=1
 /product="aminopeptidase Y"
 /protein_id="AAA19559.1"
 /db_xref="GI:469464"

/translation="MHFSLKQLAAAFYATNLGSAVYVPPQFAQFQEEPIENVLPO
 LNDSSAAVANIPIPHPIYFPMKPVSEKQDKKKDDLNATWLDRLANYSPTDI
 GHPTRVIGSGHNDKTVELINLVDDMDQYDVSLOFEALSKIIISFLNSDAETGKSF
 ANTPALSPVGVKGVKGLVPEINLCEKDYASVYPPRHNEKQIALERKGCPCFGDK
 SNLAKGFTAVYVDNPEKSGELHGLTKEPTKHTVATGVYKVGKLLATNALNI
 DYLAFAMDSYVEFTKNTIADTKHGDNDIVALGASHDSVEEGPGKNDGSGTISL
 NVKQLTHFKLNKVRFAWAAEEGLLGSNFYANLTKEENSKIRVEMDYDMWASP
 NYEVEIYDANKNEKSGEELKNLYDYIKKHLNLTLPEDGSDYVGFNNGIPAG
 GIATGAERNVNGKLVDRYCHQLCDDVSNLSWDAFITNTKLIAHSVATYADSFGFP
 KRETQKHVEDILNAQQPQFYKRADFLI"
 388..1830
 /product="aminopeptidase Y"
 /mat_peptide 707 a 474 c 434 g 657 t
 BASE COUNT
 ORIGIN

mat_peptide

alignment_scores:

Quality: 10.00 Length: 10
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-080-127-2 x YSCAMY

Align seg 1/1 to: YSCAMY from: 1 to: 2272

255 GlyproGlyLeileAsnAspGlySerGly 264

|||||

1180 GGCCGAGGTATCAACGACGATGTTCCGCT 1209

seq_name: gb_pl1:SCYBR286W

seq_documentation_block:

LOCUS SCYBR286W 2532 bp DNA PLN 11-AUG-1997
 DEFINITION S.cerevisiae chromosome II reading frame ORF YBR286W.
 ACCESSION Z36155 Y13134
 VERSION Z36155.1 GI:1870117

KEYWORDS baker's yeast.

SOURCE Saccharomyces cerevisiae

ORGANISM Saccharomycetes cerevisiae

Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;

Saccharomycetaceae; Saccharomycetes.

REFERENCE 1 (bases 1 to 2532)

Brandt, T., Christiansen, C., Holmstrom, K. and Kallensee, T.

UNPUBLISHED

REFERENCE 2 (bases 1 to 2532)

MIPS

Direct Submission

Submitted (30-AUG-1994) Data collected by MIPS on behalf of the

European yeast chromosome II sequencing project. MIPS at the

Max-Planck-Institut fuer Biochemie, Am Klopferspitze 18a D-82152

Martinsried, FRG; E-mail: Mewes@mips.emblnet.org

3 (bases 1 to 2532)

REFERENCE

AUTHORS

Barthe, C., Baur, A., Becam, A.M., Bateau, N., Boles, E., Brandt, T.,

Brendel, M., Bruckner, M., Bussereau, F., Christiansen, C.,

Contreras, R., Crouzet, M., Cziepluch, C., Demolis, N., Delaveau, T.,

Doignon, F., Domdey, H., Duesterhus, S., Dubois, E., Dujon, B., El

Bakkoury, M., Entian, K.D., Feuermann, M., Fiers, W., Fobo, G.M.,

Fritz, C., Gassenhuber, H., Glansdorff, N., Goffeau, A., Grivell, L.A.,

de Haan, M., Hein, C., Herbert, C.J., Hollenberg, C.P., Holmstrom, K.,

Jacq, C., Jacquet, M., Jauniaux, J.C., Joniaux, J.L., Kallensee, T.,

Kiesau, P., Kirchrath, L., Koetter, F., Korol, S., Liebl, S., Logghe, M.,

Lohan, A.J.E., Louis, E.J., Li, Z.Y., Maat, M.J., Mallet, L.,

Mannhaupt, G., Messenguy, F., Miosga, T., Molenans, F., Mueller, S.,

Nasr, F., Obermaier, B., Perea, J., Pierard, A., Piravandi, E.,

Pohl, F.M., Pohl, T.M., Potier, S., Proft, M., Purnelle, B., Ramezani

Scherens, B., Schwarzlose, C., Skala, J., Slonimski, P.P.,

Smits, P.H.M., Souciet, J.L., Steensma, H.Y., Stucka, R.,

Vetter, I., Vierdeels, F., Vissers, S., Wagner, G., de Weert, P.,

Wolfe, K.H., Zagulski, M., Zimmermann, F.K., Mewes, H.W. and Kleine, K.

Complete DNA sequence of yeast chromosome II

EMBO J. 13 (24), 5795-5809 (1994)

On Mar 5, 1997 this sequence version replaced gi:536735.

Location/Qualifiers

1..2532

/organism="Saccharomyces cerevisiae"

/db_xref="taxon:4932"

<1..318

/note="ORF YBR285W"

/codon_start=1

/protein_id="CAA85250.1"

/db_xref="GI:1870118"

/db_xref="SPTREMBL:P89508"

/translation="SLESPTDDSMEDISHREMATILTRNNKGVKVAIDKKGVNNHC

IDLNLKKGKLVANEHKLNDNDNRHQNTYSPEDSFEDFDKQSRILKCGSTRSYLR

CDs

```

gene      ykk"
          662..2275
          /gene="APE3"
CDS       662..2275
          /gene="APE3"
          /note="ORF YBR286w"
          /codon_start=1
          /protein_id="CA85251.1"
          /db_xref="GI:1870119"
          /db_xref="SWISS-PROT:P37302"
          /translation="MFSLKQLAVAFATNLSAVIIPQIFAQOEPIENTLPQ
          LNDSSAAVANIIPKPIPKPHESEKLDKIKTVDLNTAMDLYLAVSPDY
          GHPYVIGSKGNKMEYTLNVDMDQVYDLSIOEFLSGKITSPNLSDETKSF
          ANTTAFALSPYDGEYKGLVEIPNLGCEKDYASVPPRHNIKQIALIRGKCPFGDK
          SNLAGFGEFVAVIITDNEPKSEGLGTEPTKHTVATVGVPYKGLKLANLANT
          DYSLFAMDSYVEFKTONIADTKGDPDNTVALGSDSYEEGGLNDGSGTISL
          LNVAKOLTFKINNKYRFAMMAAEEGLGNSFYAVNLTKENSKIRPMYDMMASP
          NYEYIYDANNKNEPKSEBELNLYVDYKKAHLNLTLPVPIGRSDYVGFINGGIPAG
          GIATGAKERNVNNKYLDRCYHOICDDVSNLSMDAFITNTKLIANSVATYADSPGFP
          KRETOGKHKVYDILNMQOPQFKRADFLIT"

BASE COUNT      829 a      517 c      482 g      704 t

ORIGIN

alignment_scores:
      Quality:      10.00      Length:      10
      Ratio:      1.000      Gaps:      0
Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-080-127-2 x SCYBR286w ..
Align seq 1/1 to: SCYBR286w from: 1 to: 2532
      255 G1yProG1yleasnaSPsglySerG1y 264
      ||||||||||||||||||||||||||||
      1622 GCCCAGGTATCAGCAGATGTCGCGT 1651

seq_name: gb_p11:SCDPB3

seq_documentation_block:
LOCUS      SCDPB3      32421 bp      DNA      PLN      05-DEC-1994
DEFINITION      S.cerevisiae (s288c) Rf1, DPB3, YML27 and SNFS genes.
ACCESSION      X76053
VERSION      X76053.1 GI:600025
KEYWORDS      DPB3 gene; Rf1 gene; SNFS gene; YML27 gene.
SOURCE      baker's yeast.
ORGANISM      Saccharomyces cerevisiae
              Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
              Saccharomycetaceae; Saccharomyces.
REFERENCE      1 (bases 1 to 32420)
AUTHORS      Holmstrom,K.
TITLE      Direct Submission
JOURNAL      Submitted (05-NOV-1993) K. Holmstrom, Biotechnological Institute,
              Lundtoftevej 100, Building 227, 2800 Lyngby, DENMARK
REFERENCE      2 (bases 1 to 32420)
AUTHORS      Holmstrom,K., Brandt,T. and Kalliesoe,T.
TITLE      The sequence of a 32 420 bp segment located on the right arm of
              chromosome II from Saccharomyces cerevisiae
              yeast 10, 47-62 (1994)
              3 (bases 1 to 32421)
              Holmstrom,K.
JOURNAL      Direct Submission
REFERENCE      Submitted (05-DEC-1994) K. Holmstrom, Biotechnological Institute,
              Lundtoftevej 100, Building 227, 2800 Lyngby, DENMARK
COMMENT      On Dec 7, 1994 this sequence version replaced gi:429119.
FEATURES             Location/Qualifiers
             source          1..32421
             organism="Saccharomyces cerevisiae"
             strain="S288c"
             db_xref="taxon:4932"
             complement(1..3289)
             gene="Rf1"
             /gene="Rf1"

CDS       CDS
          /db_xref="SGD:S0000479"
          complement(1..3289)
          /gene="Rf1"
          /note="YBR1743-ORF"
          /codon_start=1
          /db_xref="SGD:S0000479"
          /protein_id="CA53638.1"
          /db_xref="GI:429120"
          /db_xref="SWISS-PROT:P29539"
          /translation="MSKDSKKTITRIDQHILRRSQHDNTSNGSSPMWKTNLPP
          SPOAMHVOISDLSPPKRRKLASSSDCKNKFDSALIKNNYKPEDTGSRMLQSVPLS
          ASNSNMVQISLPSPDRKSSPLYPYRPGSSPKRSPSSKSGKILRRRLRSVRVSD
          LSNKLOTHOKHLNNGTSTPYKTRTPNPAKSTGSEKHEGLVDNVSSEPRFIE
          GGLILRESDSEYVARRREYATRNITPIITKRVNVDOKFNLLYNINISIEICI
          PHLOIADOTLSSSEKRPVATRLVQVRFSSAISMVFKVYKTRKPDKAKVI
          IRTWGTALRNENSKNIIITAVVSFDRDKFQTFELSNELIPIISTFELIEINSHL
          IYKELLRGEFLSKYKPKMITVTSWMLGEVLPRIIDEDYSMKILTSIVLLELI
          KCLDFVEHEHRIYOCIMLSPVCEIIPKFSKLPILNDSANLDKVTIGHLTPOIK
          NTVYKNNKTIAMDLMLSMTGLDSDGKRYVDLSESNKWFEDLNLCFIIINHPYRL
          MSIKVRRITTCICIKISQKNOGKSKLILRDPQKTELPYDNDPAREGILTLG
          VYIAFISKNLSTDMELFMDHLITPIEDYERYDSIHQNTLPYLLHLLGKNA
          DVALERKKYKHIHPMSVIASEGVLKDISLPPQIKREYDKIMKVEQTEVALSNV
          NLAHDLITSLKHLPEDRKQTHLESFESLILKYQNNKDRPIFDFEGAVTSFVYT
          FLDFELRKNDLSLVNFIQISKVGISQGNMTLIDLVYIRKARNETSEFLIEFEL
          DDKTEYVAAOMWGSTLLPNTISPEREFSQILANVKKVNSIENFIDLCLKSPVN
          LFTLLHVSMSNNNTIYFIOSYVSKNEKLVADLTTLTKTSLSPDEPESGLPPLR
          NKFMIDTEYCISHNPNLUNSIPDLNSDILLKLPKRSSTFRANKIKLEKGEQTLVR
          WLKGOQLEOLNONSEIENVLQNASDSELEKSEIIRLELHAMANPIEPFSGILNF
          CIKNNMAD"
          complement(3804..6227)
          /note="YBR2013-ORF"
          /codon_start=1
          /protein_id="CA53639.1"
          /db_xref="GI:429121"
          /db_xref="SWISS-PROT:P38148"
          /translation="VLEVPSTIPGELHDLRLHODAEPECKKMFPAWDISFGQPP
          DEPHSLAIVKSDSANDNSALLRNSLEVNDIQSKMVPRTSFREGDTCTGDSNGFOY
          PNTKEILNLKFOIROLELOVDVDALENATYCNHNSILPEIKVDPGSLSEIKRYS
          RRVGSNTTLKRSQDVGRRGSLFRPPOCAKAMEMDNTVITYCSRGSGTDMOTES
          LAPCSHGNCPCVTTLLIOLICLMEVQKGVSGSLYKTNLFCTIYQNTDIPOT
          LIGPLDNEFFKNPTPLNCSPEIYCFNNVKNLMLYKELCNLTATRIEETG
          LICGNTIDMHWYQIILKNNISILTRHEENETIVNLSKSLDNDTTSISOLYNPNT
          KEVMKILIKCISNSQSPSLTKIRYLDLIDDDASRGOEHLTFEPAGSGIGLNTLI
          OSVELLNVCYLIRVOYSOVELTFMFCVEDGYTESLITFAYIIFHEHNPQDALLRI
          HRPPEFLPSPDQILIGHQPIVLRERSPENGSNKIKYANALFKRPSQLTSELSFS
          IFMKIPLSENFVNLKGPLPSILKHLTLGSLDHQNPALKSLGITHIVYGEVYSN
          TLNKDKIAHPVPRRAITMTNTNEVAGVTCNKSNRADTVVSCQENGSNVVISENS
          GFOICQIENLDNGCDPLFHOIDKVLDEISNSEATGKGLVLCMWGVSRSAIVCIAC
          MYLQCDLASAVLFVRVRLNVTIOPNFYELFKMKKKHYNREKXTMDWHIICRG
          IAEVNMKKT"
          complement(6399..6800)
          /note="YBR2014-ORF"
          /codon_start=1
          /protein_id="CA53640.1"
          /db_xref="GI:429122"
          /db_xref="SWISS-PROT:P38350"
          /translation="MFSSLFSLASFFSCMAASSRNKLSTLFSHSIEFRRLRLRY
          FPFESCAKTSDDSTFCTNNSAANAADITLVDVITYSGSHLALFFLIDIGKGAFS
          LTKLIDIFATRIALLACFYLLGOAVVKAINF"
          6478..7083
          /gene="DPB3"
          /db_xref="SGD:S0000482"
          /db_xref="SGD:S0000482"
          /db_xref="DPB3"
          /note="DPB3"
          /db_xref="YBR2015-ORF"
          /note="YBR2015-ORF"
          /codon_start=1
          /db_xref="SGD:S0000482"
          /protein_id="CA53641.1"
          /db_xref="GI:429123"
          /db_xref="SWISS-PROT:P27344"
          /db_xref="MSN:YKKAAPVFPFSLSKYKIAKCDPEYVTSNVAISATAFAEL
          FVONLVEESLIVLAQNSKGTSLPSTLSISECVKRNPRFLDEAIIQOLKNSLADK
          KRELNMGPGRSDQEVIEPELHEDDGVEEDEDEVEVEEPEYVHNEELLDDSKQON

```

```

CDS
      7441..8778
      /note="YBR2016-ORF"
      /codon_start=1
      /protein_id="CA53642.1"
      /db_xref="GI:429124"
      /db_xref="SWISS-PROT:P38351"
      /translation="MSKQVEYIAPIKYONSUPLVLPOLPPKLLVPESEPPTNDSGLIN
      SLIKTNVNLQDEDEJGMPYDLKPFGLKLDKSLKLLGIDNVDKLDRLILNDP
      RIDRLTKTIDISKVTEFLRTEYVSNTIAADNTLSKRRRLDGDSDENLVNHLISR
      VEGFNKDKQHPYKQKWKWDLPDASMDQVYFLDKFMSASLDYREKSLD
      TGFEPNLEDEEMISYATDKDSAILLENLEKMDDDSHGKRIKKRIKRIYD
      MKOYKEMPTELAIRLNDKDIAYKPLRSKTELRERRANDIKRPLEVHIDOLNVT
      LRNPSTKANTRDKLRMKDPDPIAFATVDEDEDEDEPEDEPKKESEGSKEGSGES
      ENKDEDEIKOEKENQDEENKODENRAADTPETSDAVTEQKPEBEKTELOEE"
      /complement(115127..16599)
      /note="YBR2017-ORF"
      /codon_start=1
      /protein_id="CA53643.1"
      /db_xref="GI:429125"
      /db_xref="SWISS-PROT:P38352"
      /translation="WEKIMSEVESREKEPEPAGSLPDIVQATLPPLSSDDIKNLSQIN
      KOEYTLDFDSKILMELFKAFGLTKTNDPEQGRNABFKCTETILREAPPLS
      MOEYQLRAYDAKFSWGLKGRGLYTAASNNELVATSLNGPSRFRKYVNTPEVP
      WENSRRTSRTNFTSEDPISAIRKDGDEIAOVSGSGFQILTESGNTLSGSGTFS
      GILGPGSGSOHOVNPREMIMMERISYPRITSSNGSTVNTGTGEGRMSGSHPS
      TAYEPGNATTOHITIDSGAPASPGSGSHGVPRTTPSPGPHNIYISQIEMERSA
      NKAYPGNNHIRMFRANSPPLYSRDENLGSPNDQIYAVSGSHFLAMTDNNTIS
      WDSTESDGVKIRFANLPSRATNPILKJASNMNCCYIVYGLVAMERADAIKGES
      FAFKAYEIVPNTDNGDSRIYDFACLDQDNCVFFLNNGDKLMKHNGLNDIVLNIY
      GKLCINACFASLVFTDTHCYTLKTVNGVDYKDSLELDINENYISVASGDYFVAL
      TERHVGWGTGALITIKKH"
      /complement(11154..13790)
      /note="YBR2018-ORF"
      /codon_start=1
      /protein_id="CA53644.1"
      /db_xref="GI:429126"
      /db_xref="SWISS-PROT:P38149"
      /translation="MYDSRGVALHSELHRRNNAFSLISYAFKKRLIFASODSKT
      LVDPPTNLHTTRLGSEDEHTRRSVLCVLTGSEDENFLESGADLVNIMIGEK
      TIRDFLEPTEIATVSVTIGDIFLSTAYLDSLEIYFGCNAALYENLIQIIEKK
      SSDDEVENINKLPHRYDYKFDPLSGITGSSNSLSQTSLOENCGAIIIEVPEENI
      KYAHGFYSINKLCPRENOLKESRSGAEHLIISAGDISKLMESKRGONVYK
      ISLNDKIDNEDSVISQRIEPPLYCGILDTIKIMDNTQOIIISLTKEKESVYIS
      SYMDHVAIDESCITHFYQOVNHNANQOQKIISSSEFSKSNAGSVSLTSGSDSL
      TLMDITLSLAVPLSSNPINASSSTLOTTLNMAATQASLNNEELNLTRELISQYV
      SOSKDTNTLTSLRCALYLOQLFLKFGATNSQLPLPDGPNVVFAYFQNGKSYQV
      GAKKRIIMYGHVYVLSGNTFNMTNDFLTLCENGYLKRGVADNKGPLVAISHVA
      YLFOGGLVNDVYFLESGEISGASLQVCEKEDIKIDIMLLNSNSTWVAOHP
      CLNYGLRGVINAQIKWMSDKPDGSHGGVYDPMVNLVYVSKLQNEONEIMYIP
      YSPKLDLTEEYOROKRTTELADENTYODITNTMKPLSLSTYKFSQGVITVY
      PKSVYMGISIRLVPQSYEYKRDILKALESFQOKSONLEIKVLNABRGWIGDPT
      NHAYOILKDEITIAMDEPLLIREGGSISCLRMERLFDAPAVOIPCGQSTDNHLAN
      ENLRKMSNLTJELSKVFNL"
      /db_xref="SGD:S0000486"
      /db_xref="YML27"
      /db_xref="YML27"
      /gene="YML27"
      /note="YBR2019-ORF"
      /codon_start=1
      /db_xref="SGD:S0000486"
      /protein_id="CA53645.1"
      /db_xref="GI:429127"
      /db_xref="SWISS-PROT:P36526"
      /translation="WKGSPISQFSKTSINALTIRPMKXYRDELFGYSKVGKRVPLT
      TKQGNKTMVKGTRASGIGRHTKFGGYVNMKVRVYTPDPAVNFELKPYVANVPLK
      HEFGSGGGLDPRLOLKIKEYIVNGVSEGATLDISCYERG"
      /complement(115127..16599)
      /note="YBR2020-ORF"
      /codon_start=1
      /protein_id="CA53646.1"
      /db_xref="GI:429128"
      /db_xref="SWISS-PROT:P38353"
      /translation="MSGFRIDIVKPLPLPEVELEPEFKLPEDDKITVYITAGLYL
      FAQPLVGLKRAITPNVNDPIYELRGVGEPEPRLLEFGLFPIISGLIQLAGLKV
      IKVNFKIOSDRELEFOSLTKVFAIYQVYLINIFAGYFEDDLSVOIGLINFOLVA
      GIPTTLAEVDKGEFGSGAMINVTNINVADEFGVSOIKVGDODTEOGALI
      NLIOGLSKRHKFTFGIISAFNBDYLPNTITTIYALAIIVCYLOSREYELRISR
      ARGNNVYPIKLITVGLSYLSTIIFYIHITAFVITQVANKRPHITCKIMGHE
      NANNLVPTFPLSLAPPTSFEGVYQOPLTITTSATFLVIGVIMADWQAISSG
      ARDVALFEKQOGLITLMGRQONAKELNKVPIAAVYAGASVLLIYVIGSLDKGRA
      AGIYVGIAGSGFLEVTIIEYQOOGSGSALNOVLGYVGAM"
      /db_xref="YBR2021-ORF"
      /codon_start=1
      /protein_id="CA53647.1"
      /db_xref="GI:429129"
      /db_xref="SWISS-PROT:P38150"
      /translation="MVQNMSEVFEECDYSKESPTSPRILDDLDGNDAYSQGLAFD
      GDVGITQARVRNRAQIFEDSNVDVYLHLDLDMVPLNKKPMOMEMGSPMAPAET
      PPVPEPKTKDIAVSSLAHPSEYFQTHRIDRKLCELSKIRRYNTISKDALSC
      POLHSRAGCYLPRVKEKLEGRHFDLESNTVNLQDGNVTELFNSGINTPTRE
      FREDFEWCLNITIDRSLSRSEKRLQYLVNKFVFOHLHSKEEMROSKRPHKDFYNC
      alignment_scores:
          quality: 10.00      length: 10
          ratio: 1.000      gaps: 0
          percent similarity: 100.000      percent identity: 100.000
      alignment_block:
          US-09-080-127-2 x SCDPB3
          Align seg 1/1 to: SCDPB3 from: 1 to: 32421
          255 GYPTROGLYTLAASNPASpelySergly 264
          |||||||
          21844 GCCCAAGGTATCAACGACGATGTTCCGGT 21873
      seq_name: gb_ro:AF039707
      seq_documentation_block:
          LOCUS AF039707 1428 bp mRNA 16-JAN-1998
          DEFINITION Rattus norvegicus glutamate carboxypeptidase II mRNA, partial cds.
          ACCESSION AF039707
          VERSION AF039707.1 GI:273248
          KEYWORDS
          SOURCE
          ORGANISM
          Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
          Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
          1 (bases 1 to 1428)
          Carter, R.E., Feldman, A.R. and Coyle, J.T.
          Prostate-specific membrane antigen is a hydrolase with substrate
          and pharmacologic characteristics of a neutropeptidase
          Proc. Natl. Acad. Sci. U.S.A. 93 (2), 749-753 (1996)
          96149377
          JOURNAL
          MEDLINE
          2 (bases 1 to 1428)
          Luthi-Carter, R., Berger, U.V., Barczak, A.K., Enna, M. and Coyle, J.T.
          Isolation and expression of a rat brain cDNA encoding NAALDase
          (glutamate carboxypeptidase II)
          Unpublished
          3 (bases 1 to 1428)
          Carter, R.E., Feldman, A.R. and Coyle, J.T.
          Direct Submission
          Submitted (23-DEC-1997) Laboratory of Molecular and Developmental
          Neuroscience, Mass. Gen. Hospital East, Bldg. 149, 13th St., Rm.
          2510, Charlestown, MA 02129, USA
          Location/Qualifiers
          1..1428
          /organism="Rattus norvegicus"
          /db_xref="taxon:10116"
          /clone="W6"
          /tissue_type="brain"
          /dev_stage="adult"

```

CDS
1. 1410
/EC_number="3.4.17.21"
/function="N-acetylated alpha-linked acidic dipeptidase (Naladase); N-acetylasepartlyglutamate peptidase (NAAG peptidase)"
/note="similar to prostate-specific membrane antigen (PSM, PSMa)"
/codon_start=1
/product="glutamate carboxypeptidase II"
/protein_id="AA896759.1"
/db_xref="GI:2773249"
/translation="EETEAAGLPSIPVPHVIGYDAOKLLEHMGSSAPPDSMWGGLKV
PYWPGFAGNFSKOKYKLIHSNKYRTIYNYGCTKQAEVDPRIYITGGCHDAWY
GGIDPDSGAHVHEIVRTFTGLKKGRPRRTIIFASDADAEFGELSTMAEBSHL
LOERGAVINADSSIEGNVTLRVDCPLHSLVNLKELPSDEGEGLSLDSWKE
KSPSTEIGMPRIKSLGSDNFEVFORLIGIASGRARYETNNKYSYPLHSYVE
TYELVEKFDPTERYHLTVAQVRGAWFELANSIVLPDQCSYAVALKKHAETIYNS
MNHPEKKAIVISFSLFSAVNNEFTDVASFNORLQDLKSNPLRLILNDOLMYLER
AFIDPLIGPGRPRHITVAPSSHNKYAGSEFGIYDALFDINNKTDSKAMREVKRO
ISIAETVQAARELREVD"
BASE COUNT 441 a 285 c 321 g 381 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-080-127-2 x AF039707 ..
Align seg 1/1 to: AF039707 from: 1 to: 1428

291 Alagluglupehglyleuenglyser 299
|||||
424 GCTGAAGAAATTGGCTTCTGTCT 450

seq_name: gb_Dr3:HSB01476

seq_documentation_block:
LOCUS HSB01476 1797 bp mRNA PRI 16-DEC-1999
DEFINITION Homo sapiens mRNA; CDNA DKFZp434M2227 (from clone DKFZp434M2227).
ACCESSION AL133614
VERSION AL133614.1 GI:6599227
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1797)
Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Gausenhuber,J. and Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (15-DEC-1999) MIPS, Am Klopferplatz 18a, D-82152 Martinsried, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by GBR (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://www.mips.biochem.mpg.de/proj/cDNA/.
location/Qualifiers
1. 1797
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434M2227"
/clone_1db="434 (synonym: htes3). Vector pSport1; host DH10B; sites NotI + SalI"

/dev_stage="adult"
/tissue_type="testis"
polyA_signal 1757..1762
polyA_site 1773
BASE COUNT 643 a 240 c 300 g 614 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-080-127-2 x HSB01476 ..
Align seg 1/1 to: HSB01476 from: 1 to: 1797

291 Alagluglupehglyleuenglyser 299
|||||
980 GCAGAGAAATTGGACTTCTGGGTTCC 1006

seq_name: gb_ro:AF040256

seq_documentation_block:
LOCUS AF040256 2248 bp mRNA ROD 24-MAR-1998
DEFINITION Rattus norvegicus glutamate carboxypeptidase II mRNA, complete cds.
ACCESSION AF040256
VERSION AF040256.1 GI:2982228
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 2348)
Luthi-Carter,R., Berger,U.V., Barczak,A.K., Enna,M. and Coyle,J.T.
TITLE Isolation and expression of a rat brain cDNA encoding glutamate carboxypeptidase II
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (6), 3215-3220 (1998)
MEDLINE 98169524
REFERENCE 2 (bases 1 to 2348)
Luthi-Carter,R., Berger,U.V., Barczak,A.K., Enna,M. and Coyle,J.T.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1997) Laboratory of Molecular and Developmental Neuroscience, Massachusetts General Hospital-East, Bldg. 149, 13th St., Rm. 2510, Charlestown, MA 02129, USA
location/Qualifiers
1. 2348
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/tissue_type="brain"
72. 2330
/EC_number="3.4.17.21"
/note="N-acetylated alpha-linked acidic dipeptidase (Naladase); NAAG peptidase; similar to Homo sapiens prostate-specific membrane antigen (PSMA)"
/product="glutamate carboxypeptidase II"
/codon_start=1
/protein_id="AA040067.1"
/db_xref="GI:2982229"
/translation="MMNAOODSDSAEALGRORWPCFAGTLVLAFTGFIIGFEGWEI
KPSDSSSYSGKQKAFLOELKAEKNIKFLYNPRLHAGLQTNFLAOLHAW
KERGLDVEHSDVILSTPNKHPYRISINDGNEIRKTSIAELSPGRIENISVY
PYISAKSPQTEPGDLYVNYARTEDFFELRYMKINCSGKIVLYAGGVFRONKYN
AQAGAKGIIISDPADYVGVKSPDQMNLEPGGVORGNVNLNAGDPLTPGVA
NEYAYHEFTLAVGLPSIPVPHVIGYDAOKLLEHMGSSAPPDSMWGGLKVPYVPG
FAGNFSKOKYKLIHSNKYRTIYNYGCTKQAEVDPRIYITGGCHDAWYGGIDP
GAHVHEIVRTFTGLKKGRPRRTIIFASDADAEFGELSTMAEBSHLQERGA
YINADSSIEGNVTLRVDCPLHSLVNLKELPSDEGEGLSLDSWKEKSPSTE
IGMPRIKSLGSDNFEVFORLIGIASGRARYETNNKYSYPLHSYVEYELVEK
FYDPTERYHLTVAQVRGAWFELANSIVLPDQCSYAVALKKHAETIYNSMHPDEM
KAYMISFDSLFSAVNNEFTDVASFNORLQDLKSNPLRLILNDOLMYLERAFIDPL
IGPGRPRHITVAPSSHNKYAGSEFGIYDALFDINNKTDSKAMREVKROISIAET

BASE COUNT 706 a 484 c 537 g 621 t
ORIGIN

alignment_scores:

Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-080-127-2 x AF040256 ..

Align seg 1/1 to: AF040256 from: 1 to: 2348

291 AlagluGluPheGlyLeuGlySer 299

1344 GCTGAGAGATTGGCTCTGTCTCT 1370

seq_name: gb_ptr:AF176574

seq_documentation_block:

LOCUS AF176574 2518 bp mRNA PRI 24-AUG-1999
DEFINITION Homo sapiens folylpoly-gamma-glutamate carboxypeptidase (FGCP)
MRNA, complete cds.

ACCESSION AF176574
VERSION AF176574.1 GI:5762481

KEYWORDS
SOURCE

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

REFERENCE 1 (bases 1 to 2518)
AUTHORS Devlin, A.M., Ling, E. H. and Halsted, C. H.

JOURNAL Unpublished
TITLE Intestinal folylpoly-gamma-glutamate carboxypeptidase

2 (bases 1 to 2518)

REFERENCE 2
AUTHORS Devlin, A.M., Ling, E. H. and Halsted, C. H.

JOURNAL Direct Submission
TITLE Submitted (09-AUG-1999). Internal Medicine, University of

FEATURES
source

1. .2518
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="small intestine"

gene
/gene="FGCP"

CDS
/gene="FGCP"

139. .2391
/product="folylpoly-gamma-glutamate carboxypeptidase"

/protein_id="AA051121.1"

/db_xref="GI:5762482"

/translation="MMNLHETDSAVATARRPRMLCAGLVLAGSFFLLGFLGFWFIK
SGNEATNITPKHMKAFDELKAEINKKFLYFTQIPHLAGEONFOLAKIOSQWKE
FGDSVELAHYVULISYPMKTPHNTIINENGNIFNTSLFEPPEPEANSDIYVP
PSAEPQGMPEBDLYVNVARTEDFFKLEDRKINKSGKIYARIGKVRGKYNNAO
LAGAGVILYSDPADYFAPGVSYDGMVLPGGVQKGNILNLNAGDPLPGYPAE
YVARGIAEAVGLPSIPVPIGIDYDAOKLLEMGASAPDSSMRGSLVYPYVNGFT
GNFSTQKVMHISTNEVRIYVPIGIDYDAOKLLEMGASAPDSSMRGSLVYPYVNGFT
ANFHEIVRSFGTLKKKGWPRRTILFASMDAEFGLLSTEAENSRLOBRGVAI
NADSIENGVTLRVDCPTLMYSLVNLTKEKSPDEGEGKSLYSMTKKSPSPFG
MPRIKSLSGNDFFVFORLGLASGRARTKMKTKWETNSGFLYHSYETELVYKTY
DPMERYHLTVAGVGVGVELANSIVLPDGDYAVVLRKADKLYISMKPQWMT
YVSFDSLESAVNNFTDVAENKQLOEIDKSNPYLIRIMNDQLFELRAFDIPGLP
DRPEYRHITVAPSSHNKYAGESFPGIYDALFDIESKVPDSKAMGEVKKQIYAAFTVO
AAAEITLSEVA"

Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-080-127-2 x AF176574 ..

Align seg 1/1 to: AF176574 from: 1 to: 2518

291 AlagluGluPheGlyLeuGlySer 299

1405 GCAGAGAGATTGGCTCTGTCTCT 1431

seq_name: gb_ptr:AF026380

seq_documentation_block:

LOCUS AF026380 2603 bp mRNA ROD 28-OCT-1997
DEFINITION Mus musculus prostate-specific membrane antigen homolog (mopsm)
MRNA, complete cds.

ACCESSION AF026380
VERSION AF026380.1 GI:2565337

KEYWORDS
SOURCE

house mouse.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

REFERENCE 1 (bases 1 to 2603)
AUTHORS Bacht, D.J. and Hescon, W.D.W.

JOURNAL Cloning, expression and chromosomal localization of MOPSM; the
TITLE Murine Orthologue of Prostate Specific Membrane Antigen

Unpublished
2 (bases 1 to 2603)

REFERENCE 2
AUTHORS Bacht, D.J. and Hescon, W.D.W.

JOURNAL Direct Submission
TITLE Submitted (17-SEP-1997) Surgery, Memorial Sloan-Kettering Cancer
Center, 1275 York Ave., New York, NY 10021, USA

FEATURES
source

1. .2603
/organism="Mus musculus"
/strain="Swiss Webster Nude"

/db_xref="taxon:10090"

/chromosome="7"

/map="7D1-D2"

1. .2603
/gene="mopsm"

88. .2346
/gene="mopsm"

/function="glutamate-preferring neurocarboxypeptidase;
novel folate hydrolase"

/note="human prostate-specific membrane antigen ortholog;
MOPSM; NAALDase; similar to rat NAAC-peptidase"

/codon_start=1
/product="prostate-specific membrane antigen homolog"

/protein_id="AAB81971.1"

/db_xref="GI:2565338"

/translation="MMNLHETDSAVATARRPRMLCAGLVLAGSFFLLGFLGFWFIK
SGNEATNITPKHMKAFDELKAEINKKFLYFTQIPHLAGEONFOLAKIOSQWKE
FGDSVELAHYVULISYPMKTPHNTIINENGNIFNTSLFEPPEPEANSDIYVP
PSAEPQGMPEBDLYVNVARTEDFFKLEDRKINKSGKIYARIGKVRGKYNNAO
LAGAGVILYSDPADYFAPGVSYDGMVLPGGVQKGNILNLNAGDPLPGYPAE
YVARGIAEAVGLPSIPVPIGIDYDAOKLLEMGASAPDSSMRGSLVYPYVNGFT
GNFSTQKVMHISTNEVRIYVPIGIDYDAOKLLEMGASAPDSSMRGSLVYPYVNGFT
ANFHEIVRSFGTLKKKGWPRRTILFASMDAEFGLLSTEAENSRLOBRGVAI
NADSIENGVTLRVDCPTLMYSLVNLTKEKSPDEGEGKSLYSMTKKSPSPFG
MPRIKSLSGNDFFVFORLGLASGRARTKMKTKWETNSGFLYHSYETELVYKTY
DPMERYHLTVAGVGVGVELANSIVLPDGDYAVVLRKADKLYISMKPQWMT
YVSFDSLESAVNNFTDVAENKQLOEIDKSNPYLIRIMNDQLFELRAFDIPGLP
DRPEYRHITVAPSSHNKYAGESFPGIYDALFDIESKVPDSKAMGEVKKQIYAAFTVO
AAAEITLSEVA"

BASE COUNT 747 a 491 c 604 g 676 t
ORIGIN

alignment_scores:

BASE COUNT 812 a 519 c 562 g 710 t
ORIGIN

CDS

/chromosome="xiv"
complement(158..1525)
/note="ORF YNL206c"

/codon_start=1
/protein_id="CAA96106.1"
/db_xref="GI:1302207"
/db_xref="SWISS-PROT:P40161"
/translation="MSKFLDELPESTSRKIGTVAVRVPSSLEIFEELYKALNENSDRSGRHKPRIDVSDLLKTDDEISTENTITFELEGYSVSLPRKKIDLVFIYSNDGSPVITLLKGNDRLEISTYQLNKNIMASFLVPERPNLITLMTYSCEDKFESEPVVMTNKENTLNQFNKMGLDNSVTDVEKCEVEIRKQALITGKISNPEVNSLTVDDKELNSEHLQCHRGKEGELFDPDIIEFGKRPILFDASDEISTYSITLFTNALLVKDGEKYEFEIMDOTETAKIDDYKRRKMDKMSSEELKAKSKSGOATGDDOSIQQBATROMDEKKAGVSDDEENDEFEASLSLSSGESSDGADEGEAREDEEDDEEDKKGOSALNRDNSFASINGOPEQLYKREKPELEDDIPLEIDNDDDEDDDSGVEYD"

CDS

complement(1525..1947)
/note="ORF YNL205c"
/codon_start=1
/protein_id="CAA96107.1"
/db_xref="GI:1302208"
/db_xref="SWISS-PROT:P40162"
/translation="MYTNHTKTLKARLFSQRLTFENKGRMISRDSHSCITLNSCKMRPIYRPFVLVLTGFTSLQNSIRHKNICFKNNINFEYFFIPISWTPLYNDASCFEFSNIRYSALVCSNNYDKVLETLTKKPLKRYLQI"

gene

complement(1745..2647)
/gene="SPS18"
/db_xref="SGD:S0005148"
complement(1745..2647)
/gene="SPS18"
/note="ORF YNL204c"
/codon_start=1
/db_xref="SGD:S0005148"
/protein_id="CAA96108.1"
/db_xref="GI:1302209"
/db_xref="SWISS-PROT:P32572"
/translation="MRLEFNSKDMENRRLRAKKAAGNNNCFECKSYNPQVSCSTGIFICVNCANLIRGKGTNIFCVKSTMDPEKDYRAVSGNNRFGSLKNGILONGIPLRKYNDLFAKSKYRRRLANVRNSNDNRMYGNNFOQTNGATSOIDRLREISNNSNAGCAFEVLPKYLKSGNDPCRFPACISSENNIDENNYTATSTLTLEKRONDPTGITSRSKQILSDALYKSYEDKSGSVQPTLTIENQRLPNDIKSFHNEKHLHTPLHPSVFSCFTGGDILPEFVN"

BASE COUNT 769 a 588 c 473 g 978 t

ORIGIN

Alignment_scores:
Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment_block:
US-09-080-127-2 x SCYNL206C ..

Align seg 1/1 to: SCYNL206C from: 1 to: 2808

312 Asnlylslelrgleutyrleuaenphe 320
|||||
1695 AATAAATTCGTTATATTAAATTTT 1721

seq_name: gb_r0:RN075973

seq_documentation_block:
LOCUS RN075973 2899 bp mRNA ROD 05-NOV-1997
DEFINITION Rattus norvegicus NNAg-peptidase mRNA, complete cds.
ACCESSION U75973.1
VERSION U75973.1 GI:1661226
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 2899)

AUTHORS

Bzdega, T., Turli, T., Wroblewska, B., She, D., Chung, H. S., Kim, H. and Neale, J. H.

TITLE

Molecular cloning of a peptidase against N-acetylasparylglutamate from a rat hippocampal cDNA library

JOURNAL

J. Neurochem. 69 (6), 2270-2277 (1997)

MEDLINE

98041505

REFERENCE

2 (bases 1 to 2899)

AUTHORS

Bzdega, T., Turli, T., Wroblewska, B., She, D. and Neale, J. H.

TITLE

Direct Submission

JOURNAL

Submitted (23-OCT-1996) Biology, Georgetown University, 37th & O St. NW, Washington, DC 20057, USA

location/Qualifiers

1..2899

/organism="Rattus norvegicus"

/strain="Sprague Dawley"

/db_xref="taxon:10116"

/tissue_type="brain"

23..2281

/function="peptidase activity against N-acetylasparylglutamate"

/note="NALadase; prostate specific membrane antigen; PSM; PMA"

/codon_start=1
/product="NNAg-peptidase"
/protein_id="AAC53423.1"
/db_xref="GI:1661227"

/translation="MNNAODSDSAEALGRORPCAGTLVLAFTGTFIIGPLGWFIKPNDSTSSYSYSPGKKAFIOELKAENIKFELNFRTRPLATQNFELAQINHQMKFGLDLVELSDYDLSTPNKTHPNITSLINDGNEIFETSLAESPEGEYISDVPPYSAPSPQGTPEGDLVYVYARTEDFEKLEPYKINCGRIVLARYGOVFRGNKYNAQLAGAGILYSGDPADYFVPGYKSPDGNLGGGVORONYNLNAGDPLPGVPA NEYAYRHEFEAVGLPSIPVPIGYDDAQLKLEHMGSAAPPDSMKGGLKVPYNGVGFAGNFSKOKVKLHISYNKTRYIYIGTLKGAVERDRIYLGSHRDVAVFGIDPDS GAAYVHEIYRPGTILKKKGRRPRRTILFASWDAEEGLGSTRMAEHSRLDQRCVA YINDSISBENITLRVDCITPLMSLTVYNUKLEPSPDEGBGSLDSMKREKSPSTIEIGMRISKLDSGNDVEFFQRLGIASGRARYTNMKNVSSPLTHSYETELVEK KAPYPTFHYHLTAQVAGAMAEELANSIVLPFDCOSYAVALKKAETIYINISNHPDEM FVNISDSISFVANNFTDVAASFNRLOLDLDSNPILRLINDQALYERAEIDPLGLPRPFYRHHIIVYPSHMKYAGSPFGIYDALPDIINNKYDTSKAMREVRQSIAAFTVOAAETLRREV"

BASE COUNT 896 a 573 c 612 g 818 t

ORIGIN

Alignment_scores:
Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment_block:
US-09-080-127-2 x RN075973 ..

Align seg 1/1 to: RN075973 from: 1 to: 2899

291 AlaGluglupheglyleuenglyser 299
|||||
1295 GCTGAAGAATTGGCCTTCTGTCTT 1321

seq_name: gb_pr2:HSN012370

seq_documentation_block:
LOCUS HSN012370 3152 bp mRNA PRI 25-MAR-1999
DEFINITION Homo sapiens mRNA for NALADase II protein.
ACCESSION AF012370
VERSION AF012370.1 GI:4539524
KEYWORDS
SOURCE NALADase II gene.
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3152)
Pangalos, M. N., Neelis, J. M., Somers, M., Verhasselt, P., Bekkers, M., van der Helm, L., Fraiponts, E., Ashton, D. and Gordon, R. D.

AUTHORS Chan, E.K., Imai, H., Hamel, J.C. and Tan, E.M.
TITLE Human autoantibody to RNA polymerase I transcription factor hUBF.
JOURNAL Molecular identity of nucleolus organizer region autoantigen NOR-90
MEDLINE and ribosomal RNA transcription upstream binding factor
REFERENCE 92044316
AUTHORS 2 (bases 1 to 7032)
TITLE Matera, A.G., Wu, W., Imai, H., O'Keefe, C.L. and Chan, E.K.
JOURNAL Molecular cloning of the RNA polymerase I transcription factor
MEDLINE hUBF/NOR-90 (UBTF) gene and localization to 17q21.3 by fluorescence
REFERENCE in situ hybridization and radiation hybrid mapping
AUTHORS Genomics 41 (1), 135-138 (1997)
JOURNAL 97271570
TITLE 3 (bases 1 to 7032)
AUTHORS Chan, E.K.L.
REFERENCE Direct Submission
TITLE Submitted (29-JUL-1996) Molecular and Experimental Medicine, The
JOURNAL Scripps Research Institute, 10550 North Torrey Pines Road, La
Jolla, CA 92037, USA
COMMENT On Apr 2, 1997 this sequence version replaced gi:452079.
FEATURES
source
Location/Qualifiers
1..7032
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="plasmid clone gNOR4.4"
/map="17q21.3; NOR53/33 STS maps 3.77 cR (LOD >3.0) below
framework marker UTR-9641, which maps 334.34 cR from the
top of the Chr17 linkage group"
101n(<608..672,952..1072,1430..1540,1878..2011,2137..2278,
2553..2594,2732..2845,2937..3092,3411..3566,3650..3760,
4152..4240,4344..4533,5965..6012,6106..6177,6279..7032)
/gene="UBTF"
/product="ribosomal RNA upstream binding transcription
factor"
608..7032
/gene="UBTF"
101n(<608..672,952..1072,1430..1540,1878..2011,2137..2278,
2553..2594,2732..2845,2937..3092,3411..3566,3650..3760,
4152..4240,4344..4533,5965..6012,6106..6177,6279..6692)
/gene="UBTF"
/codon_start=1
/product="ribosomal RNA upstream binding transcription
factor"
/protein_id="AAC51239.1"
/db_xref="GI:1916615"
/translation="MKYIQDFQREKQEFERNLARREDHPLLQNAKKSIDPEKPKTP
QOLWYHKKVYLKVPDATTKEVDSLGKOWSOLSKDKRLKWIHKALEORKEYEIM
RDYIOKHPELINSEEGITSTLKAERQLKDEGDRTPPPNSYSLYCAELMANMKD
VSTERMVLCSQOQKLSOKERDAYHKKCDOKKDYVELRLFLESLPEEOQVIGE
EKMLNINKKQATSPASKPAQSGGSGSEKRPVSAWFIPESEKROLOQERPELSE
SELTRLALAKMNDLSEKKAKYTKARBAALKAQSEKPGSEEREGKLPSPKRAEIIW
QOSVIGDYLAERKNDVRKALKAMETWNNKEREKLMWIKKAEDOKRYERLSEIRA
PPAATNSSKKMKFOGPKPKPNVGYOKFQSELLSNGLNHLPLKEKRWELIGSRWQIS
QSKAEHYKKLAEQOKQYKVDLWVSLSPDRAAYKEYISNKRSMYTKLRGPNPKS
SRTLOSKESEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE
GOTGSGRGESGCGHLYTPHPFSPRKKMTRTTTRMTMRKIMSPRAAAPAP
PGGTPTLTPTLEQHPRAKSPGAPLPN"
608..672
/gene="UBTF"
952..1072
/gene="UBTF"
1430..1540
/gene="UBTF"
1878..2011
/gene="UBTF"
2137..2278
/gene="UBTF"
2391..2414
/gene="UBTF"
/note="STS primer NOR53"
2553..2594
/gene="UBTF"

exon 2732..2845
/gene="UBTF"
primer_bind Complement(2803..2824)
exon /note="STS primer NOR53"
exon 2937..3092
/gene="UBTF"
exon 3411..3566
/gene="UBTF"
exon 3650..3760
/gene="UBTF"
exon 4152..4240
/gene="UBTF"
exon 4344..4533
/gene="UBTF"
exon 5965..6012
/gene="UBTF"
exon 6106..6177
/gene="UBTF"
exon 6279..7032
/gene="UBTF"
BASE COUNT 1523 a 1995 c 2190 g 1324 t
ORIGIN
Alignment_scores:
Quality: 9.00
Ratio: 1.000
Percent Similarity: 100.000
Percent Identity: 100.000
Alignment_block:
US-09-080-127-2 x HSU65487 ..
Align seg 1/1 to: HSU65487 from: 1 to: 7032
14 LeuAlaGlyArgAlaLeuValSerPro 22
|||||
2639 TTGGCAGCAGACGCCCTGTGACGCC 2665

Thu May 11 10:03:01 2000

us-09-080-127-2.rge

Page 17

N_Geneseq_36:V88034	-	7.00	92.31	1.5e+03	4441	CITTA (class II transactivator	N_Geneseq_36:V19635	+	6.00	113.11	106.11	21	Pasteurella piscicida identifi
N_Geneseq_36:V28716	-	7.00	92.14	1.6e+03	4543	Class II transactivator gene.	N_Geneseq_36:V43483	+	6.00	113.11	106.11	21	Primer alg1 for P.derrugiosa
N_Geneseq_36:V68033	-	7.00	92.14	1.6e+03	4543	Class II transactivator (CITTA	N_Geneseq_36:V48124	+	6.00	112.44	115.73	23	Staphylococcus aureus entic ge
N_Geneseq_36:V68036	-	7.00	92.11	1.6e+03	4549	CITTA (class II transactivator	N_Geneseq_36:V800021	+	6.00	112.44	115.73	23	Aspergillus oryzae aminopepti
N_Geneseq_36:V68035	-	7.00	92.11	1.6e+03	4549	CITTA (class II transactivator	N_Geneseq_36:V82525	+	6.00	112.44	115.73	23	Aspergillus oryzae aminopepti
N_Geneseq_36:V68031	-	7.00	91.97	1.6e+03	4649	CITTA (class II transactivator	N_Geneseq_36:V59651	+	6.00	112.92	120.52	24	Borrelia burgdorferi capture
N_Geneseq_36:V94032	-	7.00	91.82	1.6e+03	4746	CITTA (class II transactivator	N_Geneseq_36:V65087	+	6.00	110.17	139.59	28	Serine protein kinase PCR pri
N_Geneseq_36:V94031	-	7.00	91.79	1.6e+03	4762	Human multidrug resistance-ass	N_Geneseq_36:V10507	+	6.00	110.71	144.34	29	Scarrin gene PCR primer. fung
N_Geneseq_36:V03618	-	7.00	91.60	1.7e+03	4885	Phosphoenolpyruvate carboxylas	N_Geneseq_36:V08741	+	6.00	110.46	149.08	30	Primer VI for potato UlaSORN
N_Geneseq_36:V20621	-	7.00	91.58	1.7e+03	4901	Polynucleotide sequence from v	N_Geneseq_36:V08765	+	6.00	110.46	149.08	30	Primer VI for potato UlaSORN
N_Geneseq_36:V68016	-	7.00	91.28	1.8e+03	5105	CITTA (class II transactivator	N_Geneseq_36:V13393	+	6.00	110.46	149.08	30	Creative amidinohydroclase pri
N_Geneseq_36:V04196	-	7.00	91.06	1.8e+03	5252	Trans-feruloyl soa hydratase/	N_Geneseq_36:V84457	+	6.00	110.22	153.82	31	ICAM-4 gene domain 3', PCR p
N_Geneseq_36:V97226	-	7.00	90.91	1.8e+03	5362	Mouse additional sex comb poly	N_Geneseq_36:V11681	+	6.00	110.22	153.82	31	Human ICAM-4/ICG fusion prote
N_Geneseq_36:V97224	-	7.00	90.91	1.8e+03	5362	Mouse additional sex comb poly	N_Geneseq_36:V13467	+	6.00	110.22	153.82	31	Human ICAM-4 cDNA amplifying
N_Geneseq_36:V29323	+	7.00	90.81	1.9e+03	5438	Human calcium channel 27980/3.	N_Geneseq_36:V08895	+	6.00	110.22	153.82	31	Primer for human ICAM-4 codin
N_Geneseq_36:V68017	-	7.00	90.77	1.9e+03	5463	CITTA (class II transactivator	N_Geneseq_36:V59124	+	6.00	109.98	158.54	32	Forward PCR primer used to ge
N_Geneseq_36:V68048	-	7.00	90.77	1.9e+03	5463	CITTA (class II transactivator	N_Geneseq_36:V59124	+	6.00	109.98	158.54	32	Aspergillus oryzae aminopepti
N_Geneseq_36:V85389	+	7.00	90.08	2.0e+03	6000	Human protein tyrosine phospho	N_Geneseq_36:V800024	+	6.00	109.98	158.54	32	Aspergillus oryzae aminopepti
N_Geneseq_36:V68035	+	7.00	90.08	2.0e+03	6000	Human protein tyrosine phospho	N_Geneseq_36:V800012	+	6.00	109.98	158.54	32	Aspergillus oryzae aminopepti
N_Geneseq_36:V13352	+	7.00	89.73	2.1e+03	6285	Enterococcus faecalis genome c	N_Geneseq_36:V00012	+	6.00	109.98	158.54	32	Aspergillus oryzae aminopepti
N_Geneseq_36:V41889	+	7.00	88.45	2.2e+03	6529	Nucleotide sequence of the chl	N_Geneseq_36:V27513	+	6.00	109.32	172.06	35	Staphylococcus pneumoniae ORF
N_Geneseq_36:V62906	+	7.00	88.24	2.2e+03	7676	Human galactokinase genomic DN	N_Geneseq_36:V044793	+	6.00	109.11	177.38	36	Murine noggin probe #1. Noggi
N_Geneseq_36:V90558	-	7.00	87.97	2.7e+03	7960	Ovine Y-chromosomal DNA repeat	N_Geneseq_36:V199111	+	6.00	109.11	177.38	36	ECR1 PCR antisense primer fir
N_Geneseq_36:V35524	-	7.00	87.61	2.8e+03	8355	Human interferon-beta gene 5'	N_Geneseq_36:V04477	+	6.00	109.11	177.38	36	HCV-C linker regions of PAP-2
N_Geneseq_36:V022489	-	7.00	87.59	2.8e+03	8387	Human U266 lambda gene. B cell	N_Geneseq_36:V098451	+	6.00	108.90	182.08	37	Antisense probe DM 8. Detecti
N_Geneseq_36:V16333	+	7.00	87.58	2.8e+03	8391	MLL gene 8.3 kb fragment encod	N_Geneseq_36:V090739	+	6.00	108.90	182.08	37	Primer for human placental bl
N_Geneseq_36:V053470	+	7.00	87.58	2.8e+03	8392	MLL gene 8.3 kb BamHI genomic	N_Geneseq_36:V175150	+	6.00	108.51	191.45	39	KE108A28 for heavy chain cons
N_Geneseq_36:V37080	+	7.00	87.24	2.9e+03	8791	Human T cell Inositol 1,4,5-tri	N_Geneseq_36:V116712	+	6.00	108.32	196.12	40	3'-Primer for rat metabotropi
N_Geneseq_36:V26535	-	7.00	87.23	2.9e+03	8802	Caline von Willebrand Factor 4	N_Geneseq_36:V163041	+	6.00	107.61	214.78	44	PCR primer D for constructio
N_Geneseq_36:V022491	-	7.00	87.00	3.0e+03	9071	Human U266 lambda gene. B cell	N_Geneseq_36:V166301	+	6.00	107.45	219.43	45	Primer DBL3 used in vector p
N_Geneseq_36:V23347	-	7.00	86.81	3.1e+03	9310	U266-Lambda gene and downstre	N_Geneseq_36:V194592	+	6.00	107.45	219.43	45	Construction of trispecific p
N_Geneseq_36:V74307	-	7.00	86.75	3.1e+03	9390	Staphylococcus aureus confli	N_Geneseq_36:V15751	+	6.00	107.12	228.71	47	KE105A28 for heavy chain cons
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new Hepa	N_Geneseq_36:V99838	+	6.00	106.81	237.98	49	Human Tiam-1 gene antisense p
N_Geneseq_36:V23001	+	7.00	86.75	3.1e+03	9391	Fragment HGU1668 of a new							

N_Geneseq_36:T712133	-	6.00	96.47	896.54	197	Human acidic fibroblast growth	N_Geneseq_36:X161936	-	6.00	92.78	1.4e+03	324	Serine protease inhibitor di
N_Geneseq_36:T21213	+	6.00	96.40	905.21	199	Human histallelic polymorphic DNA	N_Geneseq_36:X161936	-	6.00	92.78	1.4e+03	324	Serine protease inhibitor di
N_Geneseq_36:T12213	+	6.00	96.18	931.21	205	Human gene signature HUMG503904	N_Geneseq_36:T060596	-	6.00	92.75	1.4e+03	325	Serine protease expressed sequen
N_Geneseq_36:T18392	+	6.00	96.07	944.19	208	Human chondrocyte RNA derived c	N_Geneseq_36:T312126	-	6.00	92.71	1.5e+03	327	E. coli j96 pathogenicity is
N_Geneseq_36:T139194	+	6.00	96.07	944.19	208	Acromiolum sp. 285 RNA-seq	N_Geneseq_36:T039727	-	6.00	92.66	1.5e+03	329	Expressed sequence tag human
N_Geneseq_36:T05341	+	6.00	96.07	944.19	208	Scopolitopsis brevicaulis 285	N_Geneseq_36:T30454	+	6.00	92.64	1.5e+03	330	Pertussis toxin B-oligomer S
N_Geneseq_36:X05341	+	6.00	96.07	944.19	208	Scopolitopsis brevicaulis spec	N_Geneseq_36:T49197	+	6.00	92.64	1.5e+03	330	Human B lymphocyte antigen B
N_Geneseq_36:X05384	+	6.00	96.07	944.19	208	Acromiolum species specific reg	N_Geneseq_36:T171924	+	6.00	92.64	1.5e+03	330	Cassette encoding human Pf4-
N_Geneseq_36:X02173	+	6.00	96.00	952.85	210	D. melanogaster chitin synthase	N_Geneseq_36:T086538	+	6.00	92.64	1.5e+03	330	EST clone AW189. New polynuc
N_Geneseq_36:X05114	+	6.00	95.96	957.17	211	Candida tropicalis 285 RNA-seq	N_Geneseq_36:T059139	+	6.00	92.62	1.5e+03	331	Human brain Expressed sequen
N_Geneseq_36:X05114	+	6.00	95.96	957.17	211	Candida tropicalis specific reg	N_Geneseq_36:T061097	+	6.00	92.62	1.5e+03	331	Human brain Expressed sequen
N_Geneseq_36:T039123	-	6.00	95.93	961.50	212	Expressed Sequence Tag human g	N_Geneseq_36:T089111	+	6.00	92.59	1.5e+03	332	EST clone BZ205. New polynuc
N_Geneseq_36:T059135	-	6.00	95.89	963.82	212	Human brain Expressed Sequence	N_Geneseq_36:T096078	+	6.00	92.57	1.5e+03	332	Bovine MHC class I gene prom
N_Geneseq_36:T077476	-	6.00	95.89	963.82	213	CNS neural thread protein 1-9a	N_Geneseq_36:T167867	+	6.00	92.57	1.5e+03	333	H. pylori surface or membran
N_Geneseq_36:T27746	-	6.00	95.82	974.46	215	Neutral thread protein partial c	N_Geneseq_36:T57825	+	6.00	92.53	1.5e+03	335	Clone R2.9#2 from mutant mul
N_Geneseq_36:T039331	+	6.00	95.82	974.46	215	Expressed Sequence Tag human g	N_Geneseq_36:T086586	+	6.00	92.51	1.5e+03	336	EST clone BA210. New polynuc
N_Geneseq_36:T059243	+	6.00	95.82	974.46	215	Human brain Expressed Sequence	N_Geneseq_36:T059514	+	6.00	92.44	1.5e+03	339	Heavy chain region of 23F2G.
N_Geneseq_36:T09824	+	6.00	95.65	996.05	220	Tomato genomic DNA, DB# 277 us	N_Geneseq_36:T819306	+	6.00	92.44	1.5e+03	339	Mouse 23F2G heavy chain vari
N_Geneseq_36:T003403	-	6.00	95.49	1.0e+03	225	Synthetic gene encoding Insulin	N_Geneseq_36:T093933	+	6.00	92.42	1.5e+03	340	Rat allograft inflammatory f
N_Geneseq_36:T25454	-	6.00	95.23	1.1e+03	233	Human gene signature HUMG507651	N_Geneseq_36:T060317	-	6.00	92.40	1.5e+03	341	Human brain Expressed sequen
N_Geneseq_36:T87948	+	6.00	95.01	1.1e+03	240	EST clone BV66. New polynucleo	N_Geneseq_36:T004595	+	6.00	92.31	1.5e+03	345	Sequence encoding metal chel
N_Geneseq_36:T08734	+	6.00	94.94	1.1e+03	242	EST clone CS319. New polynucleo	N_Geneseq_36:T040459	+	6.00	92.31	1.5e+03	345	Sequence encoding metal chel
N_Geneseq_36:X11700	-	6.00	94.79	1.1e+03	247	Human histallelic polymorphic DNA	N_Geneseq_36:T068745	+	6.00	92.31	1.5e+03	345	CH255 heavy chain variable
N_Geneseq_36:X11701	-	6.00	94.79	1.1e+03	247	Human histallelic polymorphic DNA	N_Geneseq_36:T099386	+	6.00	92.31	1.5e+03	345	Rat allograft inflammatory f
N_Geneseq_36:X12865	-	6.00	94.76	1.1e+03	248	Human histallelic polymorphic DNA	N_Geneseq_36:T87925	+	6.00	92.31	1.5e+03	345	EST clone FG448. New polynuc
N_Geneseq_36:T22894	-	6.00	94.76	1.1e+03	248	Human gene signature HUMG504351	N_Geneseq_36:T061011	+	6.00	92.29	1.5e+03	346	Human brain Expressed sequen
N_Geneseq_36:T87922	-	6.00	94.76	1.1e+03	248	EST clone BG145. New polynucleo	N_Geneseq_36:T061338	+	6.00	92.20	1.5e+03	350	Rat allograft inflammatory f
N_Geneseq_36:T08990	-	6.00	94.73	1.1e+03	249	EST clone CT333. New polynucleo	N_Geneseq_36:T029470	+	6.00	92.20	1.5e+03	350	Human gene signature HUMG506
N_Geneseq_36:X10950	-	6.00	94.67	1.1e+03	251	Human histallelic polymorphic DNA	N_Geneseq_36:T124747	+	6.00	92.20	1.5e+03	350	Human gene signature HUMG506
N_Geneseq_36:X11374	-	6.00	94.67	1.1e+03	251	Human histallelic polymorphic DNA	N_Geneseq_36:T0950161	+	6.00	92.16	1.5e+03	352	EST clone DF453. New polynuc
N_Geneseq_36:X12343	-	6.00	94.67	1.1e+03	251	Human histallelic polymorphic DNA	N_Geneseq_36:T75545	+	6.00	92.16	1.5e+03	352	Staphylococcus aureus contig
N_Geneseq_36:X12470	+	6.00	94.67	1.1e+03	251	Human histallelic polymorphic DNA	N_Geneseq_36:T060526	-	6.00	92.14	1.5e+03	353	Human brain Expressed sequen
N_Geneseq_36:T077253	+	6.00	94.61	1.2e+03	267	Human genome fragment (Preferre	N_Geneseq_36:T1417057	+	6.00	92.12	1.5e+03	354	Rice glucosyl transferase (G
N_Geneseq_36:T11501	+	6.00	94.21	1.2e+03	267	Homo sapiens adult brain clone	N_Geneseq_36:T044496	+	6.00	92.08	1.5e+03	356	Murine Noggin clone. New pol
N_Geneseq_36:T08734	+	6.00	94.19	1.2e+03	268	EST clone E293. New polynucleo	N_Geneseq_36:T1050325	+	6.00	92.05	1.5e+03	357	DK26 heavy chain variable re
N_Geneseq_36:T04048	+	6.00	94.19	1.2e+03	271	Modified kringle domain (Krl) 4	N_Geneseq_36:T53325	+	6.00	92.03	1.5e+03	358	DNA encoding a flagellar hoo
N_Geneseq_36:T010407	-	6.00	94.08	1.2e+03	272	5' region of human Retinoic Acid	N_Geneseq_36:T408514	+	6.00	92.01	1.5e+03	359	Serine protease nfp591-359 g
N_Geneseq_36:T050400	-	6.00	94.08	1.2e+03	272	HRAR-gamma D clone CDNA. New c	N_Geneseq_36:T046414	+	6.00	92.01	1.5e+03	359	Fla serine protease DNA seq
N_Geneseq_36:T24257	+	6.00	94.05	1.2e+03	273	Human gene signature HUMG507072	N_Geneseq_36:T52123	-	6.00	92.01	1.5e+03	359	Homo sapiens CBST gene relat
N_Geneseq_36:T40839	+	6.00	94.02	1.2e+03	274	Human secreted protein 5' EST S	N_Geneseq_36:T061904	-	6.00	92.01	1.5e+03	359	EST clone AK647. New polynuc
N_Geneseq_36:T09125	-	6.00	93.89	1.2e+03	279	3' nucleotide sequence of clone	N_Geneseq_36:T021197	+	6.00	91.99	1.5e+03	360	CDNA for heavy chain variabl
N_Geneseq_36:T15136	-	6.00	93.89	1.2e+03	281	Human adult testes CDNA clone A	N_Geneseq_36:T046439	+	6.00	91.99	1.5e+03	360	Mus musculus A717 antibody h
N_Geneseq_36:T086515	+	6.00	93.83	1.3e+03	281	EST clone AS330. New polynucleo	N_Geneseq_36:T029148	+	6.00	91.97	1.5e+03	361	WNI 222-5 antibody Heavy cha
N_Geneseq_36:T24218	+	6.00	93.75	1.3e+03	284	Human gene signature HUMG506789	N_Geneseq_36:T089828	-	6.00	91.93	1.5e+03	363	EST clone Hw697. New polynuc
N_Geneseq_36:T5494	+	6.00	93.73	1.3e+03	285	Minimral CMV promoter. New tetr	N_Geneseq_36:T119958	+	6.00	91.87	1.5e+03	366	Human gene signature HUMG501
N_Geneseq_36:T060892	+	6.00	93.50	1.3e+03	294	Human brain Expressed Sequence	N_Geneseq_36:T191958	+	6.00	91.85	1.5e+03	367	EST clone A2261. New polynuc
N_Geneseq_36:T125738	+	6.00	93.47	1.3e+03	295	Human gene signature HUMG507993	N_Geneseq_36:T191958	+	6.00	91.83	1.5e+03	368	Human gene signature HUMG500
N_Geneseq_36:T192851	+	6.00	93.47	1.3e+03	295	Fla saliva protein fspH(L) cDN	N_Geneseq_36:T191958	+	6.00	91.81	1.5e+03	369	Human T lymphocyte receptor
N_Geneseq_36:T125738	+	6.00	93.47	1.3e+03	295	Fla saliva protein fspH(L) cDN	N_Geneseq_36:T054521	+	6.00	91.81	1.5e+03	369	Heavy chain variable domain
N_Geneseq_36:T13399	+	6.00	93.42	1.3e+03	297	Flax gene. New rhizobial ferr	N_Geneseq_36:T078512	-	6.00	91.79	1.5e+03	370	Staphylococcus aureus contig
N_Geneseq_36:T14457	+	6.00	93.40	1.3e+03	298	Mycobacterium species nucleic a	N_Geneseq_36:T088765	+	6.00	91.71	1.5e+03	374	EST clone Hw51. New polynuc
N_Geneseq_36:T834085	+	6.00	93.37	1.3e+03	299	Mycobacterium species nucleic a	N_Geneseq_36:T1210365	+	6.00	91.71	1.5e+03	374	Polynucleotide sequence from
N_Geneseq_36:T334084	-	6.00	93.35	1.3e+03	300	Mycobacterium species nucleic a	N_Geneseq_36:T1210365	+	6.00	91.67	1.5e+03	376	3'-flanking region of VNTR 1
N_Geneseq_36:T334083	-	6.00	93.35	1.3e+03	300	Mycobacterium species nucleic a	N_Geneseq_36:T1210365	+	6.00	91.65	1.5e+03	377	Human gene signature HUMG500
N_Geneseq_36:T52120	-	6.00	93.30	1.3e+03	302	Homo sapiens CBST gene relat	N_Geneseq_36:T119515	-	6.00	91.65	1.5e+03	378	Probe (114) for microbial ge
N_Geneseq_36:T52120	-	6.00	93.30	1.3e+03	302	Human secreted protein 5' EST S	N_Geneseq_36:T02913	+	6.00	91.61	1.5e+03	378	BanH1-Sect fragment containi
N_Geneseq_36:T020302	-	6.00	93.20	1.4e+03	306	B cell hybridoma 3.i.39 VH gene	N_Geneseq_36:T020555	+	6.00	91.61	1.5e+03	379	Human brain Expressed sequen
N_Geneseq_36:T40964	-	6.00	93.18	1.4e+03	307	Human secreted protein 5' EST S	N_Geneseq_36:T060505	+	6.00	91.61	1.5e+03	379	Human brain Expressed sequen
N_Geneseq_36:T020558	+	6.00	93.13	1.4e+03	309	PAK131 encoding lipoprotein s	N_Geneseq_36:T060505	+	6.00	91.59	1.5e+03	380	Human gene signature HUMG502
N_Geneseq_36:T020558	+	6.00	93.13	1.4e+03	310	PAK131 encoding lipoprotein s	N_Geneseq_36:T121104	+	6.00	91.59	1.5e+03	380	Papillomavirus major capsid
N_Geneseq_36:T043328	+	6.00	93.10	1.4e+03	310	Tumour proliferation inhibitor	N_Geneseq_36:T49543	+	6.00	91.55	1.5e+03	380	Papillomavirus main capsid p
N_Geneseq_36:T054442	-	6.00	93.10	1.4e+03	313	P-1 cancer cell growth inhibi	N_Geneseq_36:T165439	+	6.00	91.55	1.5e+03	382	EST clone FX317. New polynuc
N_Geneseq_36:T456493	+	6.00	93.03	1.4e+03	313	Human Flap DNA. Apoptosis-inhib	N_Geneseq_36:T088377	+	6.00	91.55	1.5e+03	382	EST clone BD438. New polynuc
N_Geneseq_36:T87702	+	6.00	92.98	1.4e+03	315	EST clone EH167. New polynucleo	N_Geneseq_36:T178176	+	6.00	91.51	1.5e+03	384	EST clone EL15. New polynuc
N_Geneseq_36:T88333	+	6.00	92.96	1.4e+03	316	EST clone PER24. New polynucleo	N_Geneseq_36:T188381	+	6.00	91.47	1.5e+03	386	Papillomavirus major capsid
N_Geneseq_36:T13399	+	6.00	92.91	1.4e+03	319	Murine metastatic nucleic acid	N_Geneseq_36:T185598	+	6.00	91.42	1.5e+03	389	EST clone AS294. New polynuc
N_Geneseq_36:T169305	+	6.00	92.89	1.4e+03	321	Recombinant copolymer 1-19, am	N_Geneseq_36:T086513	+	6.00	91.40	1.5e+03	390	H. pylori cytoplasmic protei
N_Geneseq_36:T050568	+	6.00	92.84	1.4e+03	321	Rat allograft inflammatory fact	N_Geneseq_36:T49559	+	6.00	91.38	1.5e+03	393	Beet Cyst Nematode Resistanc
N_Geneseq_36:T090983	+	6.00	92.84	1.4e+03	321	Synthetic sequence capable of d	N_Geneseq_36:T160281	+	6.00	91.34	1.5e+03	393	Staphylococcal accessory reg
N_Geneseq_36:T60464	-	6.00	92.78	1.4e+03	324	Synthetic sequence capable of d	N_Geneseq_36:T060564	+	6.00	91.27	1.5e+03	397	Clone HMC546R related to Fc
N_Geneseq_36:T020303	-	6.00	92.78	1.4e+03	324	B cell hybridoma 1.5.24 VH gene	N_Geneseq_36:T44335	-	6.00	91.21	1.5e+03	400	Human brain Expressed sequen
N_Geneseq_36:T28847	-	6.00	92.78	1.4e+03	324	Desulfotribio vulgaris Mly2ak1	N_Geneseq_36:T059918	-	6.00	91.21	1.5e+03	400	Staphylococcus aureus contig
N_Geneseq_36:T90665	-	6.00	92.78	1.4e+03	324	Nucleotide sequence of clone YJ	N_Geneseq_36:T78076	-	6.00	91.21	1.5e+03	400	Staphylococcus aureus contig

N_Geneseq_36:V77896	+	6.00	91.21	1.8e+03	400	Staphylococcus aureus contig SE	N_Geneseq_36:TX16271	-	6.00	90.17	2.0e+03	460	DNA sequence of ompA-STp1. N
N_Geneseq_36:V77899	+	6.00	91.21	1.8e+03	400	Staphylococcus aureus contig SE	N_Geneseq_36:TX5983	-	6.00	90.14	2.0e+03	462	Human endothelial cell growth
N_Geneseq_36:V78585	-	6.00	91.21	1.8e+03	400	Staphylococcus aureus contig SE	N_Geneseq_36:TX22418	-	6.00	90.11	2.0e+03	464	Human gene signature HUGS04
N_Geneseq_36:TX21429	+	6.00	91.15	1.8e+03	402	M. annua pollen allergen MapR01	N_Geneseq_36:Q00564	+	6.00	90.09	2.0e+03	465	Recombinant copolymer 1-77,
N_Geneseq_36:V85126	+	6.00	91.15	1.8e+03	403	EST clone CR506. New polynucle	N_Geneseq_36:V90146	+	6.00	90.06	2.0e+03	467	EST clone DF218. New polynuc
N_Geneseq_36:V851928	+	6.00	91.15	1.8e+03	403	EST clone BA134. New polynucle	N_Geneseq_36:TX33392	+	6.00	90.06	2.0e+03	467	Human Delta7-sterol reductas
N_Geneseq_36:TX51928	+	6.00	91.15	1.8e+03	403	Human secreted protein 5' EST S	N_Geneseq_36:NT0653	+	6.00	90.04	2.0e+03	468	Fixd gene promoter. Gene for
N_Geneseq_36:Q51911	+	6.00	91.12	1.8e+03	405	Human brain Expressed Sequence	N_Geneseq_36:NT50039	+	6.00	90.04	2.0e+03	468	Sequence of the promoter of
N_Geneseq_36:TX4984	-	6.00	91.12	1.8e+03	405	Human endothelial cell growth	N_Geneseq_36:TX72727	+	6.00	90.04	2.0e+03	468	Soluble bone morphogenetic p
N_Geneseq_36:V24232	-	6.00	91.01	1.8e+03	411	Chimeric antibody against hPcRf	N_Geneseq_36:TX28024	+	6.00	90.03	2.0e+03	469	Mouse soluble BMP receptor k
N_Geneseq_36:TX00092	-	6.00	91.01	1.8e+03	411	Mouse humanised antibody #23-57	N_Geneseq_36:TX76779	-	6.00	89.95	2.1e+03	474	H. pylori transmembrane prot
N_Geneseq_36:Q060172	+	6.00	90.99	1.8e+03	412	Human brain Expressed Sequence	N_Geneseq_36:TX74459	-	6.00	89.95	2.1e+03	474	H. pylori transmembrane prot
N_Geneseq_36:Q68192	+	6.00	90.97	1.8e+03	413	Partial bovine LIF genomic clon	N_Geneseq_36:V9504	-	6.00	89.93	2.1e+03	474	Human SM73-like protein (HSM
N_Geneseq_36:TX94306	+	6.00	90.97	1.8e+03	413	PMADS gene promoter. Regulati	N_Geneseq_36:V86072	-	6.00	89.93	2.1e+03	475	EST clone d68. New polynuc
N_Geneseq_36:TX38262	+	6.00	90.95	1.8e+03	414	105 gene differentially express	N_Geneseq_36:NT01005	+	6.00	89.90	2.1e+03	477	Sequence encoding bovine ins
N_Geneseq_36:TX43442	+	6.00	90.95	1.8e+03	414	Human secreted protein 5' EST S	N_Geneseq_36:TX14824	+	6.00	89.87	2.1e+03	479	Immunoglobulin E membrane re
N_Geneseq_36:TX24502	+	6.00	90.92	1.8e+03	416	Human SR-BI gene exon 5 region.	N_Geneseq_36:TX67251	+	6.00	89.87	2.1e+03	479	Canine epsilon mb/ec segment
N_Geneseq_36:TX24537	+	6.00	90.92	1.8e+03	416	Human SR-BI gene exon 5 variant	N_Geneseq_36:V38808	+	6.00	89.85	2.1e+03	480	Homo sapiens CESP gene relat
N_Geneseq_36:TX24594	+	6.00	90.92	1.8e+03	416	Human SR-BI gene exon 5 region.	N_Geneseq_36:NT90994	-	6.00	89.84	2.1e+03	481	Partially synthetic nucleoti
N_Geneseq_36:TX24629	+	6.00	90.92	1.8e+03	417	Human SR-BI gene exon 5 variant	N_Geneseq_36:NT10029	-	6.00	89.84	2.1e+03	481	Sequence of partially synthe
N_Geneseq_36:Q38608	+	6.00	90.90	1.8e+03	417	Anti-CD4 antibody MT 15.1 heavy	N_Geneseq_36:V17755	-	6.00	89.84	2.1e+03	481	Inducible promoter (GMV'anti
N_Geneseq_36:Q43843	+	6.00	90.90	1.8e+03	417	Chimeric 128.1 VH, mouse gamma	N_Geneseq_36:V18691	-	6.00	89.84	2.1e+03	481	Cytomegalovirus immediate ea
N_Geneseq_36:TX51878	+	6.00	90.86	1.8e+03	419	Human secreted protein 5' EST S	N_Geneseq_36:TX51874	+	6.00	89.84	2.1e+03	481	Human secreted protein 5' ES
N_Geneseq_36:TX16272	+	6.00	90.85	1.8e+03	420	DNA sequence of ompA-rc-met-Slf	N_Geneseq_36:V37010	+	6.00	89.81	2.1e+03	483	Human adult placenta secreta
N_Geneseq_36:V56413	+	6.00	90.81	1.9e+03	422	Murine ICR-1.1 V-H region PCR B	N_Geneseq_36:TX91170	-	6.00	89.78	2.1e+03	484	Secreted cytokine protein-en
N_Geneseq_36:V54863	+	6.00	90.81	1.9e+03	422	Murine antibody ICR-1.1 heavy C	N_Geneseq_36:TX66477	-	6.00	89.78	2.1e+03	485	INu1 5' noncoding region. Ye
N_Geneseq_36:V61815	+	6.00	90.81	1.9e+03	422	Murine antibody ICR-1.1 Vh reg1	N_Geneseq_36:TX32599	+	6.00	89.75	2.1e+03	486	P. aeruginosa OprF C-termina
N_Geneseq_36:TX21879	+	6.00	90.79	1.9e+03	422	DNA encoding antibody against	N_Geneseq_36:TX73701	+	6.00	89.75	2.1e+03	487	Streptococcus pneumoniae pro
N_Geneseq_36:Q08193	+	6.00	90.79	1.9e+03	423	Sequence encoding C-terminal of	N_Geneseq_36:Q03871	+	6.00	89.61	2.2e+03	496	Synthetic gene encoding huma
N_Geneseq_36:Q08197	+	6.00	90.72	1.9e+03	427	Human brain Expressed Sequence	N_Geneseq_36:TX11867	+	6.00	89.58	2.2e+03	498	L. seeligeri serovar 4A nr.
N_Geneseq_36:Q53438	-	6.00	90.72	1.9e+03	427	EST locus y882e12.r1 from cDNA	N_Geneseq_36:TX07039	-	6.00	89.54	2.2e+03	501	Immunogen DNA from n-(ABCDE
N_Geneseq_36:TX93338	+	6.00	90.71	1.9e+03	428	Human brain Expressed Sequence	N_Geneseq_36:NT92034	-	6.00	89.42	2.2e+03	509	Hepatitis C virus (HCV) cDNA
N_Geneseq_36:TX47260	-	6.00	90.71	1.9e+03	428	(DSM 10101) human papillomaviru	N_Geneseq_36:NT92034	-	6.00	89.42	2.2e+03	509	Sequence of the hepatitis C
N_Geneseq_36:TX22115	+	6.00	90.65	1.9e+03	431	Human gene signature HUGS03682	N_Geneseq_36:Q28828	+	6.00	89.32	2.2e+03	516	Recombinant DNA encoding xyl
N_Geneseq_36:V90491	+	6.00	90.65	1.9e+03	431	EST clone DMV21. New polynucle	N_Geneseq_36:TX52073	+	6.00	89.32	2.2e+03	516	Nucleotide sequence encoding
N_Geneseq_36:TX41160	-	6.00	90.52	1.9e+03	439	Human secreted protein 5' EST S	N_Geneseq_36:TX75095	-	6.00	89.32	2.2e+03	516	POLYNucleotide sequence from
N_Geneseq_36:V13909	+	6.00	90.48	1.9e+03	441	Nucleotide sequence of the thir	N_Geneseq_36:V02141	-	6.00	89.27	2.3e+03	519	Human secreted protein AK647
N_Geneseq_36:TX19307	+	6.00	90.48	1.9e+03	441	cDNA sequence of the thrombopo	N_Geneseq_36:TX88061	-	6.00	89.27	2.3e+03	519	Partial cDNA clone encoding
N_Geneseq_36:Q23085	+	6.00	90.47	1.9e+03	444	Antigen mc-37c gene. Vaccine ag	N_Geneseq_36:TX13359	-	6.00	89.26	2.3e+03	520	Bidirectional promoter. New
N_Geneseq_36:Q06955	+	6.00	90.43	1.9e+03	444	Sequence encoding heavy chain	N_Geneseq_36:TX57272	-	6.00	89.26	2.3e+03	520	Bi-directional promoter. New
N_Geneseq_36:Q37721	-	6.00	90.43	1.9e+03	444	Sucrose phosphate synthase gene	N_Geneseq_36:V60079	-	6.00	89.26	2.3e+03	520	Bi-directional promoter. New
N_Geneseq_36:TX15643	+	6.00	90.43	1.9e+03	444	H. pylori ORF V3ap1820orf18. He	N_Geneseq_36:TX3907	+	6.00	89.23	2.3e+03	531	EST clone DX336. New polynuc
N_Geneseq_36:TX13540	+	6.00	90.38	2.0e+03	447	3h1 light chain variable region	N_Geneseq_36:V88235	+	6.00	89.10	2.3e+03	531	Human gene signature HUGS500
N_Geneseq_36:TX34541	+	6.00	90.38	2.0e+03	447	Monoclonal anti-idiotypic antibod	N_Geneseq_36:TX19865	+	6.00	89.08	2.3e+03	533	Human gene signature HUGS500
N_Geneseq_36:TX94344	+	6.00	90.38	2.0e+03	447	Anti-idiotypic antibody 3h1.1gh	N_Geneseq_36:TX79292	+	6.00	89.01	2.3e+03	538	Canine i9e heavy chain const
N_Geneseq_36:TX29434	+	6.00	90.38	2.0e+03	447	H. pylori cell envelope OMP ORF	N_Geneseq_36:TX00668	+	6.00	88.98	2.3e+03	540	Human secreted protein gene
N_Geneseq_36:TX72078	+	6.00	90.37	2.0e+03	448	Human gene signature HUGS02046	N_Geneseq_36:TX11110	-	6.00	88.97	2.3e+03	541	POLYNucleotide sequence from
N_Geneseq_36:V68920	+	6.00	90.35	2.0e+03	449	DNA molecule encoding a breast	N_Geneseq_36:TX20850	-	6.00	88.97	2.3e+03	541	POLYNucleotide sequence from
N_Geneseq_36:V72866	+	6.00	90.35	2.0e+03	449	Mouse 11n-54 homologue. Novel	N_Geneseq_36:TX75095	-	6.00	88.87	2.4e+03	548	Slinapis alba flowering promo
N_Geneseq_36:Q20560	+	6.00	90.33	2.0e+03	450	PAK1181 encoding lipoprotein s1	N_Geneseq_36:V68405	-	6.00	88.87	2.4e+03	548	EST clone DBJ12. New polynuc
N_Geneseq_36:TX73454	+	6.00	90.33	2.0e+03	450	PHCMW*-1 promoter. Regulatory	N_Geneseq_36:TX50757	+	6.00	88.79	2.4e+03	554	Partial sequence of ovine IL
N_Geneseq_36:Q76267	-	6.00	90.33	2.0e+03	450	PHCMW*-2 promoter. Regulatory	N_Geneseq_36:Q20556	+	6.00	88.78	2.4e+03	555	PAK191 encoding lipoprotein
N_Geneseq_36:TX06869	-	6.00	90.33	2.0e+03	450	PHCMW*-1 tecto construct. Polynu	N_Geneseq_36:TX11338	+	6.00	88.75	2.4e+03	557	AFT-1 interacting protein (p
N_Geneseq_36:TX06870	-	6.00	90.33	2.0e+03	450	PHCMW*-2 tecto construct. Polynu	N_Geneseq_36:Q20557	+	6.00	88.74	2.4e+03	558	PAK191 encoding lipoprotein
N_Geneseq_36:TX11355	-	6.00	90.33	2.0e+03	450	Minimal CMV promoter and ten te	N_Geneseq_36:V69406	+	6.00	88.72	2.4e+03	559	EHRLichia sp. HGE-7 3'-end D
N_Geneseq_36:TX11356	-	6.00	90.33	2.0e+03	450	Minimal CMV promoter and ten te	N_Geneseq_36:V90277	+	6.00	88.64	2.4e+03	565	EST clone DJ167. New polynuc
N_Geneseq_36:V60080	-	6.00	90.33	2.0e+03	450	Cytomegalovirus minimal promot	N_Geneseq_36:TX67961	-	6.00	88.62	2.4e+03	567	H. pylori transmembrane prot
N_Geneseq_36:V60081	-	6.00	90.33	2.0e+03	450	Cytomegalovirus minimal promot	N_Geneseq_36:TX30811	-	6.00	88.60	2.5e+03	568	Streptococcus pneumoniae gen
N_Geneseq_36:TX01365	-	6.00	90.33	2.0e+03	450	PHCMW*-1 sequence. Transgenic	N_Geneseq_36:TX11358	+	6.00	88.59	2.5e+03	569	Bidirectional promoter. New
N_Geneseq_36:TX01367	-	6.00	90.33	2.0e+03	450	PHCMW*-2 sequence. Transgenic	N_Geneseq_36:TX11358	+	6.00	88.59	2.5e+03	569	Bidirectional promoter. New
N_Geneseq_36:TX01421	-	6.00	90.33	2.0e+03	450	PHCMW*-1 promoter sequence. New	N_Geneseq_36:TX57121	+	6.00	88.59	2.5e+03	569	Bi-directional tetraacycline-
N_Geneseq_36:TX21901	-	6.00	90.33	2.0e+03	450	PHCMW*-2 promoter sequence. New	N_Geneseq_36:TX15721	+	6.00	88.59	2.5e+03	569	Bi-directional tetraacycline-
N_Geneseq_36:V88309	-	6.00	90.30	2.0e+03	452	EST clone DX112. New polynucle	N_Geneseq_36:V60078	+	6.00	88.59	2.5e+03	569	Bidirectional promoter regio
N_Geneseq_36:Q50606	+	6.00	90.28	2.0e+03	453	Human IL-1 chimeric antibody CI	N_Geneseq_36:V60078	+	6.00	88.59	2.5e+03	569	Bidirectional promoter regio
N_Geneseq_36:Q50609	+	6.00	90.28	2.0e+03	453	Nucleotide sequence of the 3'	N_Geneseq_36:V75297	+	6.00	88.56	2.5e+03	571	Staphylococcus aureus contig
N_Geneseq_36:V09280	-	6.00	90.28	2.0e+03	453	Staphylococcus aureus contig SE	N_Geneseq_36:Q34348	+	6.00	88.55	2.5e+03	572	Sequence of xylanase xyl 1.
N_Geneseq_36:Q02491	-	6.00	90.27	2.0e+03	454	cDNA encoding human acidic fibr	N_Geneseq_36:TX16657	+	6.00	88.54	2.5e+03	573	Hepatitis C virus isolate HK
N_Geneseq_36:Q03873	-	6.00	90.27	2.0e+03	454	Synthetic gene encoding human	N_Geneseq_36:Q83873	+	6.00	88.50	2.5e+03	576	Hepatitis C virus envelope US
N_Geneseq_36:Q01016	-	6.00	90.27	2.0e+03	454	Human acidic fibroblast growth	N_Geneseq_36:TX16587	+	6.00	88.45	2.5e+03	580	HCV protease active region.
N_Geneseq_36:TX01039	-	6.00	90.27	2.0e+03	454	Human acidic fibroblast growth	N_Geneseq_36:TX16316	+	6.00	88.45	2.5e+03	580	E. coli y36 pathogenicity is
N_Geneseq_36:Q25916	-	6.00	90.27	2.0e+03	454	ARGF mutacin. New acidic fibrobl	N_Geneseq_36:V60361	-	6.00	88.45	2.5e+03	580	EST clone B208. New polynuc
N_Geneseq_36:V60361	-	6.00	90.27	2.0e+03	454	DNA sequence of fibroblast grow	N_Geneseq_36:V60361	-	6.00	88.45	2.5e+03	580	EST clone B208. New polynuc
N_Geneseq_36:TX20063	+	6.00	90.17	2.0e+03	460	Enterococcus faecalis EFO33 gen	N_Geneseq_36:Q62079	+	6.00	88.38	2.5e+03	585	L. brevis IFO3954 16S/23S am

N_Geneseq_36:V088312	+	6.00	88.38	2.5e+03	585	Lactothacilium brevis 16S/23S rR	N_Geneseq_36:V076361	+	6.00	87.05	3.0e+03	700	Human immunoglobulin epsilon 1 rep
N_Geneseq_36:V29347	+	6.00	88.32	2.6e+03	590	Calcium ion channel alpha1 subu	N_Geneseq_36:V78780	+	6.00	87.03	3.0e+03	702	DNA encoding an aldose 1 rep
N_Geneseq_36:Q07802	+	6.00	88.31	2.6e+03	591	Human epsilon chain C-terminal	N_Geneseq_36:V837713	+	6.00	87.01	3.0e+03	704	DNA encoding a techolase 1 rep
N_Geneseq_36:V01212	+	6.00	88.31	2.6e+03	591	Human Ige Epsilon chain membran	N_Geneseq_36:V339693	+	6.00	86.97	3.0e+03	708	Gastric cancer associated g
N_Geneseq_36:V11226	+	6.00	88.28	2.6e+03	593	Porphyromonas gingivalis target	N_Geneseq_36:V211139	+	6.00	86.94	3.0e+03	710	Polynucleotide sequence fro
N_Geneseq_36:N93069	+	6.00	88.20	2.6e+03	600	Fragment of pCR-eik.2 Recombina	N_Geneseq_36:Q05711	+	6.00	86.93	3.0e+03	711	TRF61. Single polypeptide c
N_Geneseq_36:V33336	+	6.00	88.20	2.6e+03	600	DNA encoding a Staphylococcus a	N_Geneseq_36:Q51577	+	6.00	86.93	3.0e+03	711	Coding sequence of polypept
N_Geneseq_36:V17506	+	6.00	88.15	2.6e+03	604	Arabidopsis flowering promoting	N_Geneseq_36:V137376	+	6.00	86.93	3.0e+03	711	Single chain binding molecu
N_Geneseq_36:V18851	+	6.00	88.13	2.6e+03	605	Human tub genomic sequence - ex	N_Geneseq_36:V766460	+	6.00	86.93	3.0e+03	711	TRF61 single chain binding
N_Geneseq_36:V161717	+	6.00	88.13	2.6e+03	605	Human tub gene exon 4 sequence	N_Geneseq_36:V766310	+	6.00	86.90	3.1e+03	714	Single chain anti-distalopa
N_Geneseq_36:V173595	+	6.00	88.12	2.6e+03	606	Leukemia inhibiting factor gen	N_Geneseq_36:V704219	+	6.00	86.87	3.1e+03	717	Anti-EGRF single chain anti
N_Geneseq_36:V173949	+	6.00	88.12	2.6e+03	606	Obesity gene promoter (first ex	N_Geneseq_36:V782244	+	6.00	86.86	3.1e+03	718	EST clone B0358. New polyu
N_Geneseq_36:V140367	+	6.00	88.11	2.6e+03	607	Human secreted protein 5' EST S	N_Geneseq_36:V678860	+	6.00	86.83	3.1e+03	721	EPV 42 kD promoter/Human IL
N_Geneseq_36:V203038	+	6.00	88.10	2.6e+03	608	Borrelia burgdorferi polyuococ	N_Geneseq_36:V347717	+	6.00	86.83	3.1e+03	721	C. albicans antigenic prote
N_Geneseq_36:V183736	+	6.00	88.06	2.6e+03	611	DNA encoding a Staphylococcus a	N_Geneseq_36:V320701	+	6.00	86.82	3.1e+03	722	Polynucleotide sequence fro
N_Geneseq_36:V141897	+	6.00	88.05	2.6e+03	612	Poetal myelin basic protein MBP	N_Geneseq_36:V710429	+	6.00	86.66	3.2e+03	738	Sequence encoding human Xa1
N_Geneseq_36:V11897	+	6.00	88.05	2.6e+03	612	Poetal myelin basic protein MBP	N_Geneseq_36:V020293	+	6.00	86.66	3.2e+03	738	Human B7-2 extracellular do
N_Geneseq_36:V178033	+	6.00	88.04	2.6e+03	613	Mouse incomplete BMP type II re	N_Geneseq_36:Q06182	+	6.00	86.64	3.2e+03	740	PRB125 partial sequence co
N_Geneseq_36:V178933	+	6.00	88.01	2.7e+03	615	Murine BCL-XL/BCL-2 associated	N_Geneseq_36:Q057139	+	6.00	86.63	3.2e+03	741	TRF59. Single polypeptide c
N_Geneseq_36:V278934	+	6.00	88.01	2.7e+03	615	Mutant BCL-XL/BCL-2 associated	N_Geneseq_36:Q51558	+	6.00	86.63	3.2e+03	741	Coding sequence of polypept
N_Geneseq_36:V27835	+	6.00	88.01	2.7e+03	615	Mutant BCL-XL/BCL-2 associated	N_Geneseq_36:V137671	+	6.00	86.63	3.2e+03	741	Single chain binding molecu
N_Geneseq_36:V27836	+	6.00	88.01	2.7e+03	615	Mutant BCL-XL/BCL-2 associated	N_Geneseq_36:V326431	+	6.00	86.63	3.2e+03	741	TRF59 single chain binding
N_Geneseq_36:Q211769	+	6.00	88.00	2.7e+03	616	Sequence encoding plasmodium fa	N_Geneseq_36:V86872	+	6.00	86.62	3.2e+03	742	Ripening banana pulp DNA c
N_Geneseq_36:Q53551	+	6.00	88.00	2.7e+03	616	HCYV (Towme) major immediate ea	N_Geneseq_36:V891735	+	6.00	86.59	3.2e+03	745	Cloned Vitreoscilla haemoglob
N_Geneseq_36:V777193	+	6.00	88.00	2.7e+03	616	HCYV immediate early enhancer 8	N_Geneseq_36:Q118972	+	6.00	86.59	3.2e+03	745	Vitreoscilla promoter/regula
N_Geneseq_36:V052023	+	6.00	87.96	2.7e+03	619	Murine retinolic acid receptor 8	N_Geneseq_36:Q34236	+	6.00	86.59	3.2e+03	745	Vitreoscilla haemoglobin ge
N_Geneseq_36:V201130	+	6.00	87.96	2.7e+03	619	Polynucleotide sequence from th	N_Geneseq_36:V339732	+	6.00	86.59	3.2e+03	745	Gastric cancer associated g
N_Geneseq_36:V201130	+	6.00	87.96	2.7e+03	619	Polynucleotide sequence from th	N_Geneseq_36:V339732	+	6.00	86.59	3.2e+03	745	Gastric cancer associated g
N_Geneseq_36:Q076745	+	6.00	87.92	2.7e+03	623	Caenorhabditis elegans her-1 ge	N_Geneseq_36:V62573	+	6.00	86.58	3.2e+03	746	scfV(t.1)MSM). gene. T cells
N_Geneseq_36:V076706	+	6.00	87.90	2.7e+03	624	Human Bax protein cDNA. Metho	N_Geneseq_36:V100027	+	6.00	86.57	3.2e+03	747	Prostate cancer associated
N_Geneseq_36:V8969012	+	6.00	87.90	2.7e+03	624	Human alpha meltrin cDNA fram	N_Geneseq_36:Q057139	+	6.00	86.51	3.2e+03	753	TRF104b. Single polypeptide
N_Geneseq_36:V896990	+	6.00	87.90	2.7e+03	624	EST clone CW07. New polynucle	N_Geneseq_36:Q51559	+	6.00	86.51	3.2e+03	753	Coding sequence of polypept
N_Geneseq_36:V84005	+	6.00	87.90	2.7e+03	624	CDNA encoding a human BCL-2 ass	N_Geneseq_36:V137378	+	6.00	86.51	3.2e+03	753	Single chain binding molecu
N_Geneseq_36:Q02962	+	6.00	87.88	2.7e+03	626	Human IL-1ra BAC contigunous DNA	N_Geneseq_36:V336462	+	6.00	86.51	3.2e+03	753	TRF104b single chain bindin
N_Geneseq_36:V191468	+	6.00	87.80	2.7e+03	633	Mycobacterium tuberculosis anti	N_Geneseq_36:V317447	+	6.00	86.48	3.2e+03	756	Malate isomerase coding seq
N_Geneseq_36:V191405	+	6.00	87.80	2.7e+03	633	Mycobacterium tuberculosis anti	N_Geneseq_36:V317447	+	6.00	86.48	3.2e+03	756	Malate isomerase coding seq
N_Geneseq_36:V143434	+	6.00	87.80	2.7e+03	633	Mycobacterium tuberculosis anti	N_Geneseq_36:V334748	+	6.00	86.48	3.2e+03	756	DNA encoding p28 protein. N
N_Geneseq_36:V444452	+	6.00	87.80	2.7e+03	633	M. tuberculosis immunogenic pol	N_Geneseq_36:V872729	+	6.00	86.47	3.2e+03	757	Human c-myc exon 2 fragment
N_Geneseq_36:V433025	+	6.00	87.77	2.7e+03	635	Multiple sclerosis associated r	N_Geneseq_36:V25137	+	6.00	86.45	3.2e+03	759	H. pylori cytoplasmic prote
N_Geneseq_36:V329702	+	6.00	87.77	2.7e+03	635	Clone CL6-3' from MSRV-1. Nucle	N_Geneseq_36:Q92526	+	6.00	86.44	3.2e+03	760	P. communis (pear) arabinog
N_Geneseq_36:V309813	+	6.00	87.77	2.7e+03	635	Gastric cancer associated gene.	N_Geneseq_36:V3393205	+	6.00	86.42	3.3e+03	762	Interleukin (IL)-12A cDNA
N_Geneseq_36:V159717	+	6.00	87.75	2.7e+03	637	hRAR-gamma A clone cDNA. New cD	N_Geneseq_36:Q057139	+	6.00	86.41	3.3e+03	763	Glatiracetam-S-transferase (
N_Geneseq_36:N93088	+	6.00	87.74	2.7e+03	638	Acidic fibroblast growth factor	N_Geneseq_36:V3398451	+	6.00	86.37	3.3e+03	767	Gastric cancer associated g
N_Geneseq_36:V107288	+	6.00	87.74	2.7e+03	638	Complete cDNA sequence of human	N_Geneseq_36:V734361	+	6.00	86.35	3.3e+03	770	ATM gene exons 21-22. New I
N_Geneseq_36:V173503	+	6.00	87.74	2.7e+03	638	Human beta-endothelial cell gro	N_Geneseq_36:V90653	+	6.00	86.34	3.3e+03	770	Nucleotide sequence of clon
N_Geneseq_36:V343330	+	6.00	87.74	2.7e+03	638	Human endothelial cell growth f	N_Geneseq_36:V133314	+	6.00	86.33	3.3e+03	771	OSPA variant #3 coding sequ
N_Geneseq_36:V101733	+	6.00	87.74	2.7e+03	638	Human endothelial cell growth f	N_Geneseq_36:V133314	+	6.00	86.31	3.3e+03	773	Sequence of the 5' flankin
N_Geneseq_36:V159595	+	6.00	87.73	2.8e+03	639	Human endothelial cell growth f	N_Geneseq_36:V121220	+	6.00	86.31	3.3e+03	773	Sequence of the 5' flankin
N_Geneseq_36:V1395757	+	6.00	87.73	2.8e+03	639	Gastric cancer associated gene	N_Geneseq_36:V10212	+	6.00	86.31	3.3e+03	773	Sequence of 5' flankin reg
N_Geneseq_36:Q07887	+	6.00	87.66	2.8e+03	645	CDNA contg. an ORF for a globul	N_Geneseq_36:V10211	+	6.00	86.31	3.3e+03	773	Sequence of 5' flankin reg
N_Geneseq_36:V132600	+	6.00	87.66	2.8e+03	645	P. aeruginosa oprf-optR hybrid	N_Geneseq_36:V133313	+	6.00	86.30	3.3e+03	774	OSPA variant #1 coding sequ
N_Geneseq_36:V463311	+	6.00	87.62	2.8e+03	648	Human secreted protein cDNA fro	N_Geneseq_36:V133313	+	6.00	86.30	3.3e+03	774	OSPA variant #2 coding sequ
N_Geneseq_36:Q54930	+	6.00	87.54	2.8e+03	655	T. nivaum cyclophilin gene frag	N_Geneseq_36:V3398524	+	6.00	86.30	3.3e+03	774	Prostate tumour specific g
N_Geneseq_36:V159230	+	6.00	87.51	2.8e+03	658	Nucleotide sequence of the spec	N_Geneseq_36:V611169	+	6.00	86.30	3.3e+03	774	CDNA sequence of prostate t
N_Geneseq_36:Q097485	+	6.00	87.49	2.8e+03	660	Human interleukin-12 35 kd subu	N_Geneseq_36:V44861	+	6.00	86.29	3.3e+03	775	Clone BV66.1 coding sequenc
N_Geneseq_36:V072714	+	6.00	87.49	2.8e+03	660	DNA encoding human interleukin-	N_Geneseq_36:Q25793	+	6.00	86.27	3.3e+03	777	Encodes modified soluble Bo
N_Geneseq_36:V072714	+	6.00	87.49	2.8e+03	660	Human interleukin-12 p35 gene.	N_Geneseq_36:V133301	+	6.00	86.27	3.3e+03	777	OSPA soluble variant coding
N_Geneseq_36:V3374768	+	6.00	87.42	2.8e+03	666	Human secreted protein cDNA fro	N_Geneseq_36:V33857	+	6.00	86.27	3.3e+03	777	Plant PA enzyme DNA sequen
N_Geneseq_36:V150758	+	6.00	87.42	2.8e+03	666	Sequence of ovine IL-12 35 kd s	N_Geneseq_36:V133689	+	6.00	86.25	3.3e+03	780	ACYLP ORF 92. Residues 794/
N_Geneseq_36:Q105199	+	6.00	87.41	2.9e+03	667	5' region of Human Retinolic A	N_Geneseq_36:V69494	+	6.00	86.21	3.3e+03	784	Banana fruit ripening-relat
N_Geneseq_36:V015356	+	6.00	87.41	2.9e+03	667	hRAR-gamma C clone cDNA. New cD	N_Geneseq_36:V505087	+	6.00	86.17	3.4e+03	788	Staphylococcus aureus confi
N_Geneseq_36:V137375	+	6.00	87.32	2.9e+03	675	Nucleotide sequence of the bovi	N_Geneseq_36:V309806	+	6.00	86.15	3.4e+03	790	Plant SAR gene PST-1. New D
N_Geneseq_36:V29157	+	6.00	87.32	2.9e+03	675	Nucleotide sequence of the bovi	N_Geneseq_36:V309806	+	6.00	86.15	3.4e+03	790	Plant SAR gene PST-1. New D
N_Geneseq_36:V2X0062	+	6.00	87.32	2.9e+03	678	Enterococcus faecalis gene EF03	N_Geneseq_36:V399862	+	6.00	86.14	3.4e+03	791	Human secreted protein gene
N_Geneseq_36:Q005710	+	6.00	87.29	2.9e+03	678	Coding sequence of polypeptide	N_Geneseq_36:V57770	+	6.00	86.10	3.4e+03	795	H. pylori inner membrane pr
N_Geneseq_36:Q133755	+	6.00	87.29	2.9e+03	678	Single chain binding molecule T	N_Geneseq_36:V303302	+	6.00	86.08	3.4e+03	798	Methanococcus jennus phosm
N_Geneseq_36:V113755	+	6.00	87.29	2.9e+03	678	Single chain binding molecule T	N_Geneseq_36:V303302	+	6.00	86.07	3.4e+03	799	Enterococcus faecalis phosm
N_Geneseq_36:V140114	+	6.00	87.27	2.9e+03	680	Gastric cancer associated gene.	N_Geneseq_36:V3133748	+	6.00	86.05	3.4e+03	801	H. pylori GHO 343 gene. Ne
N_Geneseq_36:V134419	+	6.00	87.25	2.9e+03	681	P. aeruginosa oprf-optR hybrid	N_Geneseq_36:V110136	+	6.00	85.99	3.4e+03	807	Human elastase IV gene. New
N_Geneseq_36:Q25466	+	6.00	87.22	2.9e+03	684	Sequence of the genomic childr1	N_Geneseq_36:V11536	+	6.00	85.99	3.4e+03	807	DNA encoding human caldecoll
N_Geneseq_36:V133690	+	6.00	87.21	2.9e+03	685	Enterococcus faecalis genome c	N_Geneseq_36:V115450	+	6.00	85.94	3.4e+03	813	Lung cancer specific antiye
N_Geneseq_36:V112716	+	6.00	87.17	3.0e+03	689	E. coli j36 pathogenicity islan	N_Geneseq_36:Q12419	+	6.00	85.93	3.5e+03	814	Sequence from approx. -2300
N_Geneseq_36:V696919	+	6.00	87.10	3.0e+03	689	Human secreted protein gene 9 c	N_Geneseq_36:Q20706	+	6.00	85.88	3.5e+03	819	OSPA gene. Compositions for
N_Geneseq_36:Q11886	+	6.00	87.06	3.0e+03	695	Human serine protease 60 (SP60)	N_Geneseq_36:Q202013	+	6.00	85.88	3.5e+03	819	OSPA gene variant 25015. Co
N_Geneseq_36:V79127	+	6.00	87.06	3.0e+03	699	Human serine protease 60 (SP60)	N_Geneseq_36:Q78951	+	6.00	85.88	3.5e+03	819	Human immunoglobulin Vh gen

N_Geneseq_36:090712	6.00	85.88	3.5e+03	819	B31/25015 outer surface protease	N_Geneseq_36:V37402	6.00	85.05	3.9e+03	916	Streptococcus pneumoniae cod
N_Geneseq_36:090718	6.00	85.87	3.5e+03	819	B. burgdorferi strain 25015 out	N_Geneseq_36:V37402	6.00	85.03	3.9e+03	916	DNA sequence of portion of p
N_Geneseq_36:102361	6.00	85.87	3.5e+03	820	TIMP-3 clone 10mpHCA-3. New Li	N_Geneseq_36:V37402	6.00	85.03	3.9e+03	919	Sequence of a portion of the
N_Geneseq_36:102361	6.00	85.87	3.5e+03	820	Human gene fragment D86c11 from	N_Geneseq_36:V37402	6.00	85.03	3.9e+03	919	Streptococcus facialis genome
N_Geneseq_36:090709	6.00	85.87	3.5e+03	821	B. burgdorferi strain B31 outer	N_Geneseq_36:V37402	6.00	84.99	3.9e+03	921	Human IL-1 gene exon-10
N_Geneseq_36:011095	6.00	85.86	3.5e+03	822	Lyme disease 31 kD antigen OspA	N_Geneseq_36:V37402	6.00	84.98	3.9e+03	922	Streptococcus pneumoniae cod
N_Geneseq_36:081558	6.00	85.86	3.5e+03	822	DNA encoding OspA, a 31 kDa ant	N_Geneseq_36:V37402	6.00	84.97	3.9e+03	925	Streptococcus pneumoniae cod
N_Geneseq_36:081558	6.00	85.86	3.5e+03	822	DNA encoding OspA, a 31 kDa ant	N_Geneseq_36:V37402	6.00	84.97	3.9e+03	926	Wheat Type III glutathione t
N_Geneseq_36:086037	6.00	85.86	3.5e+03	822	Burkholderia burgdorferi OspB/OspB	N_Geneseq_36:V37402	6.00	84.95	3.9e+03	928	Hydrol DNA cloning a fami
N_Geneseq_36:115431	6.00	85.86	3.5e+03	822	Lung cancer specific antigen HC	N_Geneseq_36:V37402	6.00	84.94	3.9e+03	930	Enhancer for eukaryotic exp
N_Geneseq_36:090717	6.00	85.86	3.5e+03	822	B31/K48 outer surface protein A	N_Geneseq_36:V37402	6.00	84.94	3.9e+03	930	Sequence of HCMV(Ad169) prom
N_Geneseq_36:090729	6.00	85.86	3.5e+03	822	B31/K48 outer surface prote	N_Geneseq_36:V37402	6.00	84.94	3.9e+03	930	Sequence of HCMV(Ad169) prom
N_Geneseq_36:090730	6.00	85.86	3.5e+03	822	B31/K48 outer surface prote	N_Geneseq_36:V37402	6.00	84.94	3.9e+03	930	HCMV (Ad169) major enhance
N_Geneseq_36:090717	6.00	85.86	3.5e+03	822	B31 outer surface protein A and	N_Geneseq_36:V37402	6.00	84.94	3.9e+03	930	HCMV immediate early enhance
N_Geneseq_36:148481	6.00	85.86	3.5e+03	822	Bax omega protein cDNA. Poly	N_Geneseq_36:V37402	6.00	84.91	3.9e+03	933	Sequence encoding human bone
N_Geneseq_36:044853	6.00	85.84	3.5e+03	824	Mouse soluble EPO receptor prot	N_Geneseq_36:V37402	6.00	84.91	3.9e+03	934	H. pylori GHPD 1270 gene. Ne
N_Geneseq_36:038088	6.00	85.83	3.5e+03	825	ZOI Osp A DNA. DNA for sequen	N_Geneseq_36:V37402	6.00	84.87	4.0e+03	938	Homo sapiens BE123 sequence.
N_Geneseq_36:090722	6.00	85.83	3.5e+03	825	Pttdb/Ppo outer surface prote	N_Geneseq_36:V37402	6.00	84.84	4.0e+03	942	Rat CD86 coding sequence. Ra
N_Geneseq_36:090725	6.00	85.83	3.5e+03	825	K48/Bga outer surface prote	N_Geneseq_36:V37402	6.00	84.84	4.0e+03	943	Thelaria buffeli main ptop
N_Geneseq_36:090710	6.00	85.83	3.5e+03	825	B. burgdorferi strain K48 outer	N_Geneseq_36:V37402	6.00	84.83	4.0e+03	944	B. burgdorferi strain B31 Os
N_Geneseq_36:168152	6.00	85.83	3.5e+03	825	H. pylori cytoplasmic protein C	N_Geneseq_36:V37402	6.00	84.77	4.0e+03	951	18S rRNA gene. Detection of
N_Geneseq_36:V69390	6.00	85.83	3.5e+03	825	Human RLP ras-like protein DNA.	N_Geneseq_36:V37402	6.00	84.77	4.0e+03	951	H. pylori outer membrane pro
N_Geneseq_36:V74896	6.00	85.78	3.5e+03	830	Staphylococcus aureus contig SE	N_Geneseq_36:V37402	6.00	84.76	4.0e+03	953	DNA encoding alpha-1,4-gluc
N_Geneseq_36:V14171	6.00	85.78	3.5e+03	830	H. pylori GHPD 599 gene. New	N_Geneseq_36:V37402	6.00	84.75	4.0e+03	954	B. burgdorferi strain 1p90 O
N_Geneseq_36:V03230	6.00	85.78	3.5e+03	831	DNA encoding CD86 extracellular	N_Geneseq_36:V37402	6.00	84.75	4.0e+03	954	Rye chitinase-like protein C
N_Geneseq_36:V42925	6.00	85.75	3.5e+03	833	DNA encoding a human prostate	N_Geneseq_36:V37402	6.00	84.73	4.0e+03	956	STM7 (suppressor of telomere
N_Geneseq_36:V24248	6.00	85.71	3.5e+03	836	Human retinaldehyde binding pr	N_Geneseq_36:V37402	6.00	84.72	4.0e+03	958	B. burgdorferi strain B31 Os
N_Geneseq_36:V28498	6.00	85.71	3.5e+03	838	Novel disease associated membra	N_Geneseq_36:V37402	6.00	84.71	4.0e+03	959	H. pylori outer membrane pro
N_Geneseq_36:V98715	6.00	85.70	3.5e+03	840	DNA encoding a S. pneumoniae pr	N_Geneseq_36:V37402	6.00	84.71	4.0e+03	959	Human homologue of rat elast
N_Geneseq_36:V17177	6.00	85.68	3.5e+03	842	Brucella abortus 15 kDa antigen	N_Geneseq_36:V37402	6.00	84.69	4.0e+03	961	Murine leptin receptor spli
N_Geneseq_36:V07117	6.00	85.68	3.5e+03	843	Ehrlichia chaffeensis MAP1 gene	N_Geneseq_36:V37402	6.00	84.68	4.0e+03	963	TIMP-3 DNA clone (Timp3)-clone
N_Geneseq_36:V14178	6.00	85.67	3.5e+03	843	H. pylori GHPD 713 gene. New	N_Geneseq_36:V37402	6.00	84.65	4.0e+03	965	HAMPV IE-1 gene (Timp3)-clone
N_Geneseq_36:V33969	6.00	85.67	3.5e+03	846	Renal cancer associated gene. N	N_Geneseq_36:V37402	6.00	84.65	4.0e+03	967	Aspergillus oryzae tannase g
N_Geneseq_36:V11312	6.00	85.64	3.5e+03	846	Tunicate serine protease cDNA.	N_Geneseq_36:V37402	6.00	84.61	4.0e+03	972	B7-2 cDNA. Use of immunode
N_Geneseq_36:V33969	6.00	85.64	3.5e+03	846	DNA encoding OMP-1 protein. Nov	N_Geneseq_36:V37402	6.00	84.60	4.0e+03	973	Rat syntaxin 4 gene. Screen
N_Geneseq_36:V60304	6.00	85.60	3.5e+03	851	Aspergillus niger xylanase cDNA	N_Geneseq_36:V37402	6.00	84.58	4.0e+03	976	H. pylori GHPD 1555 gene. Ne
N_Geneseq_36:V09035	6.00	85.59	3.5e+03	852	T. sergentii 33 kDa protein. New	N_Geneseq_36:V37402	6.00	84.56	4.0e+03	978	Sequence encoding goat IGF-1
N_Geneseq_36:V04615	6.00	85.59	3.5e+03	852	Colon cancer associated gene. N	N_Geneseq_36:V37402	6.00	84.54	4.0e+03	981	H. pylori cytoplasmic prot
N_Geneseq_36:V04615	6.00	85.56	3.5e+03	855	Colan serine protease gene. N	N_Geneseq_36:V37402	6.00	84.54	4.0e+03	981	H. pylori cytoplasmic prot
N_Geneseq_36:V080618	6.00	85.56	3.5e+03	855	Kidney injury associated molec	N_Geneseq_36:V37402	6.00	84.50	4.0e+03	986	Lung cancer specific antigen
N_Geneseq_36:V113639	6.00	85.55	3.5e+03	855	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.50	4.0e+03	986	HCAV11 phosphotyrosin-modi
N_Geneseq_36:V01213	6.00	85.55	3.5e+03	856	Cytotoxic lymphocyte maturation	N_Geneseq_36:V37402	6.00	84.47	4.0e+03	990	Thelaria sergentii main piro
N_Geneseq_36:V177850	6.00	85.55	3.5e+03	856	35 kDa subunit of human cytotox	N_Geneseq_36:V37402	6.00	84.46	4.0e+03	990	Thermotoga maritima endoglu
N_Geneseq_36:V000402	6.00	85.55	3.5e+03	856	35 kDa subunit of human cytotox	N_Geneseq_36:V37402	6.00	84.46	4.0e+03	992	Marker DNA from blight-stufe
N_Geneseq_36:V197970	6.00	85.55	3.5e+03	856	35 kDa subunit of human cytotox	N_Geneseq_36:V37402	6.00	84.46	4.0e+03	992	Human ICM-4 clone 18 cDNA.
N_Geneseq_36:V37772	6.00	85.55	3.5e+03	856	Cytotoxic lymphocyte maturation	N_Geneseq_36:V37402	6.00	84.46	4.0e+03	992	Human ICM-4 cDNA clone 18
N_Geneseq_36:V34746	6.00	85.51	3.5e+03	861	DNA encoding OMP-1D protein. No	N_Geneseq_36:V37402	6.00	84.46	4.0e+03	992	Alternately spliced cDNA o
N_Geneseq_36:V33969	6.00	85.47	3.5e+03	866	Gastric cancer associated gene.	N_Geneseq_36:V37402	6.00	84.45	4.0e+03	993	Transglutaminase (expressed
N_Geneseq_36:V41589	6.00	85.46	3.5e+03	867	Human epimorphin coding sequen	N_Geneseq_36:V37402	6.00	84.44	4.0e+03	994	Streptococcus pneumoniae SP1
N_Geneseq_36:V075244	6.00	85.46	3.5e+03	867	Wild type human epimorphin gen	N_Geneseq_36:V37402	6.00	84.42	4.0e+03	997	Human calcium channel 27980/
N_Geneseq_36:V74674	6.00	85.46	3.5e+03	867	Staphylococcus aureus contig SE	N_Geneseq_36:V37402	6.00	84.41	4.0e+03	999	Mycobacterium tuberculosis a
N_Geneseq_36:V374028	6.00	85.43	3.5e+03	871	Salmonella secreted protein Ssp	N_Geneseq_36:V37402	6.00	84.41	4.0e+03	999	Mycobacterium tuberculosis a
N_Geneseq_36:V374028	6.00	85.43	3.5e+03	871	Salmonella secreted protein Ssp	N_Geneseq_36:V37402	6.00	84.41	4.0e+03	999	Mycobacterium tuberculosis a
N_Geneseq_36:V374028	6.00	85.43	3.5e+03	871	Salmonella secreted protein Ssp	N_Geneseq_36:V37402	6.00	84.41	4.0e+03	999	Mycobacterium tuberculosis a
N_Geneseq_36:V60049	6.00	85.42	3.5e+03	872	ompA-tc-met-secretory leukocyte	N_Geneseq_36:V37402	6.00	84.41	4.0e+03	999	Mycobacterium tuberculosis a
N_Geneseq_36:V84645	6.00	85.42	3.5e+03	872	Human secreted protein gene 18	N_Geneseq_36:V37402	6.00	84.35	4.0e+03	1006	DNA fragment functional as
N_Geneseq_36:V14415	6.00	85.40	3.5e+03	874	Sequence from Tashkent isolate	N_Geneseq_36:V37402	6.00	84.35	4.0e+03	1006	Promoter probe vector from
N_Geneseq_36:V544732	6.00	85.40	3.5e+03	874	Partial DNA sequence of the T	N_Geneseq_36:V37402	6.00	84.33	4.0e+03	1006	Promoter from B.flavum M23
N_Geneseq_36:V63634	6.00	85.40	3.5e+03	874	ET-RANB (HEV) partial Tashkent	N_Geneseq_36:V37402	6.00	84.33	4.0e+03	1009	Human secreted protein gene
N_Geneseq_36:V70094	6.00	85.38	3.5e+03	876	Sequence of region in front of	N_Geneseq_36:V37402	6.00	84.31	4.0e+03	1012	Human catalytic telomerase
N_Geneseq_36:V051818	6.00	85.37	3.5e+03	877	Nucleotide sequence of cold sh	N_Geneseq_36:V37402	6.00	84.27	4.0e+03	1017	NS1-257 fusion coding seque
N_Geneseq_36:V38313	6.00	85.35	3.5e+03	880	DNA encoding a Staphylococcus	N_Geneseq_36:V37402	6.00	84.27	4.0e+03	1017	Goat cytochrome 14 kDa prote
N_Geneseq_36:V183919	6.00	85.35	3.5e+03	880	DNA encoding a Staphylococcus	N_Geneseq_36:V37402	6.00	84.27	4.0e+03	1017	Kurthia sp. biotin syntheti
N_Geneseq_36:V81770	6.00	85.31	3.5e+03	885	Sequence encoding immunoreacti	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	NS1-Q01 fusion coding seque
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00	85.27	3.5e+03	889	Enterococcus faecalis genome co	N_Geneseq_36:V37402	6.00	84.25	4.0e+03	1020	O. sativa promoter #1. New
N_Geneseq_36:V13350	6.00										

N_Geneseq_36:T49316	-	6.00	84.14	4.4e+03	1035	Hypersensitive response elicitor
N_Geneseq_36:V39330	-	6.00	84.14	4.4e+03	1035	Pseudomonas solanaceae hyper
N_Geneseq_36:V39375	-	6.00	84.14	4.4e+03	1035	Pseudomonas solanaceae hyper
N_Geneseq_36:V39391	-	6.00	84.14	4.4e+03	1035	DNA encoding a hypersensitive
N_Geneseq_36:V39399	-	6.00	84.14	4.4e+03	1035	Secreted protein CSF19.1 cDNA
N_Geneseq_36:V39399	-	6.00	84.14	4.4e+03	1037	EBV VCA-P-40 gene ORF Bdfp1, F
N_Geneseq_36:V39399	-	6.00	84.14	4.4e+03	1038	E. coli biotin synthetase (bld
N_Geneseq_36:V39399	-	6.00	84.14	4.4e+03	1041	Bacterial transglutaminase, Pr
N_Geneseq_36:V39399	-	6.00	84.07	4.4e+03	1045	Enterococcus faecalis EF014 g
N_Geneseq_36:V39399	-	6.00	84.06	4.4e+03	1047	DNA fragment functional as a p
N_Geneseq_36:V39399	-	6.00	84.05	4.4e+03	1048	Streptococcus pneumoniae codin
N_Geneseq_36:V39399	-	6.00	84.04	4.4e+03	1049	Human secreted protein gene 48
N_Geneseq_36:V39399	-	6.00	84.04	4.4e+03	1050	Human SR calcium ion release c
N_Geneseq_36:V39399	-	6.00	84.04	4.4e+03	1050	Human SR calcium ion release c
N_Geneseq_36:V39399	-	6.00	84.04	4.4e+03	1050	WT human SR Ca2+ release chan
N_Geneseq_36:V39399	-	6.00	84.04	4.4e+03	1050	Mutant human SR Ca2+ release c
N_Geneseq_36:V39399	-	6.00	83.99	4.4e+03	1056	Insect poxvirus 38 kDa protein
N_Geneseq_36:V39399	-	6.00	83.94	4.5e+03	1064	Basic chitinase/lysozyme prote
N_Geneseq_36:V39399	-	6.00	83.94	4.5e+03	1064	Basic tobacco chitinase/lysozy
N_Geneseq_36:V39399	-	6.00	83.94	4.5e+03	1064	Tobacco chitinase gene, Isol
N_Geneseq_36:V39399	-	6.00	83.93	4.5e+03	1065	Porcine CD34 from clone IAX-3
N_Geneseq_36:V39399	-	6.00	83.93	4.5e+03	1065	Human I(3)mbt protein coding s
N_Geneseq_36:V39399	-	6.00	83.89	4.5e+03	1071	Prostaglandin D receptor codin
N_Geneseq_36:V39399	-	6.00	83.87	4.5e+03	1071	Streptothricin acetyl transfer
N_Geneseq_36:V39399	-	6.00	83.85	4.5e+03	1076	Insert from lambda RAI-1 encod
N_Geneseq_36:V39399	-	6.00	83.83	4.5e+03	1079	Alternatively spliced cDNA of
N_Geneseq_36:V39399	-	6.00	83.83	4.5e+03	1080	Human myelin oligodendrocyte g
N_Geneseq_36:V39399	-	6.00	83.83	4.5e+03	1080	Human myelin oligodendrocyte g
N_Geneseq_36:V39399	-	6.00	83.83	4.5e+03	1080	Human secreted protein gene 54
N_Geneseq_36:V39399	-	6.00	83.82	4.5e+03	1080	H. pylori GHPD 430 gene, New 1
N_Geneseq_36:V39399	-	6.00	83.82	4.5e+03	1081	Human secreted protein gene 11
N_Geneseq_36:V39399	-	6.00	83.81	4.5e+03	1082	Acidic fibroblast growth facto
N_Geneseq_36:V39399	-	6.00	83.80	4.6e+03	1084	E. coli B10 B gene, Plasmid con
N_Geneseq_36:V39399	-	6.00	83.78	4.6e+03	1087	Human endogenous retroviral DN
N_Geneseq_36:V39399	-	6.00	83.77	4.6e+03	1088	Nucleotide sequence of the chl
N_Geneseq_36:V39399	-	6.00	83.76	4.6e+03	1090	Wheat geminivirus Repr binding
N_Geneseq_36:V39399	-	6.00	83.74	4.6e+03	1092	Promoter of Aspergillus oryzae
N_Geneseq_36:V39399	-	6.00	83.72	4.6e+03	1095	C. rubicans TR1B coding sequ
N_Geneseq_36:V39399	-	6.00	83.72	4.6e+03	1095	C. rubicans TR1B coding sequ
N_Geneseq_36:V39399	-	6.00	83.72	4.6e+03	1095	Enterococcus faecalis genome c
N_Geneseq_36:V39399	-	6.00	83.70	4.6e+03	1098	Staphylococcus aureus contig S
N_Geneseq_36:V39399	-	6.00	83.69	4.6e+03	1098	Human secreted protein gene 77
N_Geneseq_36:V39399	-	6.00	83.69	4.6e+03	1100	Primate interleukin-11, New ma
N_Geneseq_36:V39399	-	6.00	83.68	4.6e+03	1101	Human secreted protein gene 13
N_Geneseq_36:V39399	-	6.00	83.67	4.6e+03	1103	Sequence encoding ovine interl
N_Geneseq_36:V39399	-	6.00	83.66	4.6e+03	1104	Lung cancer specific antigen H
N_Geneseq_36:V39399	-	6.00	83.64	4.6e+03	1107	Corynebacterium 1107 bp promoter fr
N_Geneseq_36:V39399	-	6.00	83.61	4.6e+03	1107	Promoter from B. flavum M123 R
N_Geneseq_36:V39399	-	6.00	83.61	4.6e+03	1112	Plasmodium falciparum ribonuc
N_Geneseq_36:V39399	-	6.00	83.61	4.6e+03	1112	H. pylori GHPD 655 gene, New 1
N_Geneseq_36:V39399	-	6.00	83.58	4.7e+03	1116	Murine alpha-1,3-galactosyltran
N_Geneseq_36:V39399	-	6.00	83.57	4.7e+03	1118	Murine macrophage inflammatory
N_Geneseq_36:V39399	-	6.00	83.56	4.7e+03	1119	Streptococcus pneumoniae genom
N_Geneseq_36:V39399	-	6.00	83.56	4.7e+03	1120	Human B lymphocyte antigen B7-
N_Geneseq_36:V39399	-	6.00	83.56	4.7e+03	1120	Human B lymphocyte antigen B7-
N_Geneseq_36:V39399	-	6.00	83.56	4.7e+03	1120	Human B7-2 antigen coding sequ
N_Geneseq_36:V39399	-	6.00	83.55	4.7e+03	1121	Sequence encoding biotin synth
N_Geneseq_36:V39399	-	6.00	83.54	4.7e+03	1122	MP4 chimera (MBP21.5-delta PHF
N_Geneseq_36:V39399	-	6.00	83.54	4.7e+03	1123	Humicola xylinase gene, Asperg
N_Geneseq_36:V39399	-	6.00	83.52	4.7e+03	1123	Humicola insolens xylinase, AS
N_Geneseq_36:V39399	-	6.00	83.52	4.7e+03	1125	PM4 chimera (delta PHF4-MBP21
N_Geneseq_36:V39399	-	6.00	83.50	4.7e+03	1128	Slalyltransferase coding sequ
N_Geneseq_36:V39399	-	6.00	83.50	4.7e+03	1128	Tomato endo-xyloglucan transfe
N_Geneseq_36:V39399	-	6.00	83.50	4.7e+03	1128	Tomato endo-xyloglucan transfe
N_Geneseq_36:V39399	-	6.00	83.48	4.7e+03	1132	Pinus radiata cone-specific PR
N_Geneseq_36:V39399	-	6.00	83.48	4.7e+03	1132	N. excelsior water channel pro
N_Geneseq_36:V39399	-	6.00	83.46	4.8e+03	1134	Sequence encoding peptide used
N_Geneseq_36:V39399	-	6.00	83.45	4.8e+03	1136	Human endogenous retrovirus W
N_Geneseq_36:V39399	-	6.00	83.43	4.8e+03	1139	DNA encoding a S. pneumoniae
N_Geneseq_36:V39399	-	6.00	83.42	4.8e+03	1141	Staphylococcus aureus contig S
N_Geneseq_36:V39399	-	6.00	83.42	4.8e+03	1144	Stylosanthes humilis Shp66 per
N_Geneseq_36:V39399	-	6.00	83.39	4.8e+03	1146	Human secreted protein gene 17
N_Geneseq_36:V55785	-	6.00	83.35	4.8e+03	1151	Mouse B7-2 antigen coding s
N_Geneseq_36:V55785	-	6.00	83.35	4.8e+03	1151	MP3 chimera (MBP21.5-delta
N_Geneseq_36:V55785	-	6.00	83.32	4.9e+03	1161	Human bisphosphonate bindin
N_Geneseq_36:V55785	-	6.00	83.32	4.9e+03	1163	Murine B lymphocyte antigen
N_Geneseq_36:V55785	-	6.00	83.28	4.9e+03	1163	Murine B lymphocyte antigen
N_Geneseq_36:V55785	-	6.00	83.28	4.9e+03	1163	Human endogenous retroviral
N_Geneseq_36:V55785	-	6.00	83.27	4.9e+03	1164	Complete gene sequence of w
N_Geneseq_36:V55785	-	6.00	83.27	4.9e+03	1164	Complete gene sequence of w
N_Geneseq_36:V55785	-	6.00	83.27	4.9e+03	1164	Sarcosine oxidase M gene, N
N_Geneseq_36:V55785	-	6.00	83.27	4.9e+03	1164	S. griseofuscus glucose iso
N_Geneseq_36:V55785	-	6.00	83.27	4.9e+03	1164	S. griseofuscus glucose iso
N_Geneseq_36:V55785	-	6.00	83.27	4.9e+03	1164	Enterococcus faecalis gene
N_Geneseq_36:V55785	-	6.00	83.26	4.9e+03	1165	Mouse CCR8 chemokine gene,
N_Geneseq_36:V55785	-	6.00	83.25	4.9e+03	1167	DNA sequence encoding Scrp
N_Geneseq_36:V55785	-	6.00	83.25	4.9e+03	1167	DNA fragment containing acy
N_Geneseq_36:V55785	-	6.00	83.25	4.9e+03	1167	Insert from pPRP1R1 encodin
N_Geneseq_36:V55785	-	6.00	83.24	4.9e+03	1167	Sequence encoding an intein
N_Geneseq_36:V55785	-	6.00	83.24	4.9e+03	1168	Canarypox C6 locus right fl
N_Geneseq_36:V55785	-	6.00	83.23	4.9e+03	1170	Human ribonucleotide reduct
N_Geneseq_36:V55785	-	6.00	83.22	4.9e+03	1172	Human glandular kallikrein
N_Geneseq_36:V55785	-	6.00	83.22	4.9e+03	1172	Human glandular kallikrein
N_Geneseq_36:V55785	-	6.00	83.22	4.9e+03	1172	Human glandular kallikrein
N_Geneseq_36:V55785	-	6.00	83.21	4.9e+03	1173	Cheryl spheroidin gene, Sphe
N_Geneseq_36:V55785	-	6.00	83.21	4.9e+03	1173	Single chain anti-dsialoga
N_Geneseq_36:V55785	-	6.00	83.18	4.9e+03	1178	DNA contg. repeats of tuber
N_Geneseq_36:V55785	-	6.00	83.17	4.9e+03	1180	Human gene fragment Q2944 f
N_Geneseq_36:V55785	-	6.00	83.14	5.0e+03	1185	Human secreted protein 8 en
N_Geneseq_36:V55785	-	6.00	83.11	5.0e+03	1190	Renal cancer antigen HOM-RC
N_Geneseq_36:V55785	-	6.00	83.11	5.0e+03	1190	Renal cell carcinoma antige
N_Geneseq_36:V55785	-	6.00	83.06	5.0e+03	1195	DNA encoding a g1a24 poly
N_Geneseq_36:V55785	-	6.00	83.06	5.0e+03	1198	Sequence of nm-sigma intron
N_Geneseq_36:V55785	-	6.00	83.04	5.0e+03	1201	Human retinaldehyde bindin
N_Geneseq_36:V55785	-	6.00	83.03	5.0e+03	1203	Dengue virus type 4 envelop
N_Geneseq_36:V55785	-	6.00	83.03	5.0e+03	1203	Mouse Interleukin-B30 (IL-B
N_Geneseq_36:V55785	-	6.00	83.01	5.0e+03	1205	Homo sapiens 20q13 amplicon
N_Geneseq_36:V55785	-	6.00	83.01	5.0e+03	1205	Nucleotide sequence encodin
N_Geneseq_36:V55785	-	6.00	82.99	5.1e+03	1208	snac gene encoding enzyme 1
N_Geneseq_36:V55785	-	6.00	82.99	5.1e+03	1209	Porcine acylglucosamine-2-e
N_Geneseq_36:V55785	-	6.00	82.98	5.1e+03	1210	Rat fatty acid binding prot
N_Geneseq_36:V55785	-	6.00	82.97	5.1e+03	1210	Human secreted protein gene
N_Geneseq_36:V55785	-	6.00	82.97	5.1e+03	1212	Creatinine amidinohydrolyase
N_Geneseq_36:V55785	-	6.00	82.97	5.1e+03	1212	Creatinine amidinohydrolyase
N_Geneseq_36:V55785	-	6.00	82.97	5.1e+03	1212	Human RDS4 nucleic acid se
N_Geneseq_36:V55785	-	6.00	82.97	5.1e+03	1212	Stable creatine amidinohydr
N_Geneseq_36:V55785	-	6.00	82.95	5.1e+03	1215	Creatine amidinohydrolyase,
N_Geneseq_36:V55785	-	6.00	82.95	5.1e+03	1216	CMV-PSA promoter, Nucleic a
N_Geneseq_36:V55785	-	6.00	82.86	5.1e+03	1230	Xenorhabdus luminescens fla
N_Geneseq_36:V55785	-	6.00	82.86	5.1e+03	1230	BTG-contg. sequence, Prep.
N_Geneseq_36:V55785	-	6.00	82.86	5.1e+03	1230	Sequence of the replication
N_Geneseq_36:V55785	-	6.00	82.86	5.1e+03	1236	Nucleotide sequence encodin
N_Geneseq_36:V55785	-	6.00	82.82	5.2e+03	1236	Nucleotide sequence of clus
N_Geneseq_36:V55785	-	6.00	82.82	5.2e+03	1236	DNA encoding a tafz enzyme
N_Geneseq_36:V55785	-	6.00	82.82	5.2e+03	1236	Poly nucleotide sequence firo
N_Geneseq_36:V55785	-	6.00	82.80	5.2e+03	1240	TIMP-3 DNA clone Tim3-clon
N_Geneseq_36:V55785	-	6.00	82.80	5.2e+03	1240	Prostaglandin D receptor fu
N_Geneseq_36:V55785	-	6.00	82.80	5.2e+03	1240	Mouse p32.11-cla-retinol de
N_Geneseq_36:V55785	-	6.00	82.80	5.2e+03	1240	CDNA clone encoding a retin
N_Geneseq_36:V55785	-	6.00	82.77	5.2e+03	1245	Human BNP, Purified mammali
N_Geneseq_36:V55785	-	6.00	82.77	5.2e+03	1245	5-enolpyruvylshikimate-3-ph
N_Geneseq_36:V55785	-	6.00	82.74	5.2e+03	1250	Calcium activated potassium
N_Geneseq_36:V55785	-	6.00	82.72	5.2e+03	1252	Alpha-acetolactate dehydrat
N_Geneseq_36:V55785	-	6.00	82.72	5.2e+03	1253	Outer membrane protein F of
N_Geneseq_36:V55785	-	6.00	82.72	5.2e+03	1254	Human gamma-1 chain first m
N_Geneseq_36:V55785	-	6.00	82.72	5.2e+03	1254	ATM gene exons 48-49, New 1
N_Geneseq_36:V55785	-	6.00	82.72	5.2e+03	1254	H. pylori GHPD 1072 gene, N
N_Geneseq_36:V55785	-	6.00	82.72	5.2e+03	1254	Staphylococcus aureus mutan
N_Geneseq_36:V55785	-	6.00	82.72	5.2e+03	1254	Staphylococcus aureus mutan
N_Geneseq_36:V55785	-	6.00	82.72	5.2e+03	1254	Staphylococcus aureus mutan
N_Geneseq_36:V55785	-	6.00	82.70	5.2e+03	1257	Sequence of the replication
N_Geneseq_36:V55785	-	6.00	82.70	5.2e+03	1257	H. pylori secreted or perip
N_Geneseq_36:V55785	-	6.00	82.68	5.2e+03	1257	H. pylori secreted or perip
N_Geneseq_36:V55785	-	6.00	82.68	5.2e+03	1261	Mouse B7-2 exons m1B, 2, 3,
N_Geneseq_36:V55785	-	6.00	82.66	5.3e+03	1263	Nad B gene encoding L aspar

seq_name: N_Geneseq_36:013381 + 6.00 82.52 5.4e+03 1287 Human beta2-adrenergic-STR2
N_Geneseq_36:013381 + 6.00 82.54 5.4e+03 1285 TIMP-3 metalloproteinase inh1b
N_Geneseq_36:013381 + 6.00 82.54 5.4e+03 1284 DNA encoding 5-enolpyruvylshik
N_Geneseq_36:013381 + 6.00 82.54 5.4e+03 1280 CDNA encoding ornithine carban
N_Geneseq_36:013381 + 6.00 82.54 5.4e+03 1278 Chlamydomonas reinhardtii hist
N_Geneseq_36:013381 + 6.00 82.54 5.4e+03 1276 Human tissue inhibitor of met
N_Geneseq_36:013381 + 6.00 82.54 5.4e+03 1275 Human secreted protein gene 11
N_Geneseq_36:013381 + 6.00 82.54 5.4e+03 1275 Nucleotide sequence of protein
N_Geneseq_36:013381 + 6.00 82.54 5.4e+03 1269 Maize associated region MAR2
N_Geneseq_36:013381 + 6.00 82.54 5.4e+03 1267 Homo sapiens Fyb71-81 encoding
N_Geneseq_36:013381 + 6.00 82.54 5.4e+03 1267 Duffy blood group gpd protein
N_Geneseq_36:013381 + 6.00 82.54 5.4e+03 1265 Human retinol binding protein

seq_name: N_Geneseq_36:013381 + 6.00 82.52 5.4e+03 1287

seq_documentation_block:
ID: X00070 standard; DNA: 1491 BP.

AC: X00070;

DT: 17-MAR-1999 (first entry)

DE: Aspergillus oryzae dipeptidyl aminopeptidase encoding DNA #2.

KW: Prolyl dipeptidyl aminopeptidase; protein hydrolyzate; dough;

KW: Flavour enhancer; palatability; mouthfeel; aroma; crust colour;

KW: Baking; animal feed additive; hydrolysis; ss.

OS: Aspergillus oryzae.

FT: Key

FT: CDS

FT: 1. 1491

FT: /*tag= a

FT: W09851803.A1.

FT: 19-NOV-1998.

FT: 12-MAY-1998; U09629.

FT: 20-OCT-1997; US-062892.

FT: 16-MAY-1997; US-857884.

FT: (NOVO.) NOVO NORDISK BIOTECH INC.

FT: Binkovskiy A. Brown K. Byun T. Klotz A. Key MW.

FT: WPI: 99-045232/04.

FT: P-PSDA; W89614.

FT: New dipeptidyl aminopeptidase from Aspergillus oryzae - used to

FT: produce protein hydrolysates enriched in particular amino acids,

FT: PT: useful as flavour enhancers, e.g. in doughs

FT: PS: Claim 1; Fig 1; 77pp; English

FT: The present sequence encodes dipeptidyl aminopeptidase (DPA) from

FT: Aspergillus oryzae. DPA acts synergistically with an aminopeptidase

FT: (AP) to hydrolyse polypeptides, producing protein hydrolysate (PH),

FT: CC: useful in foods as flavour enhancer, e.g. in baked goods, enriched in:

FT: CC: (a) Ala, Arg, Asp, Gly and/or Val, or (b), if the substrate has been

FT: CC: deaminated, in Glu (free and/or peptide bound), in which case products

FT: CC: are useful as animal feed additives. DPA can also be used in flavour-

FT: CC: improving compositions (optionally containing AP) and in dough pre-mixes,

FT: CC: also for deactivating enzymes and for converting precursors to mature

FT: CC: proteins. DPA increases the level of hydrolysis of proteins and thus

FT: CC: of flavour development, and a mixture with AP may hydrolyse tripeptides

FT: CC: that are resistant to either enzyme used alone. PH have improved

FT: CC: solubility, emulsifying and foaming properties, and products containing

FT: CC: them have better flavour, palatability, mouthfeel, aroma and crust

FT: CC: colour.

FT: SO: Sequence 1491 BP; 335 A; 438 C; 426 G; 292 T;

FT: Alignment_scores:

FT: Quality: 496.00 Length: 496

FT: Ratio: 1.000 Gaps: 0

FT: Percent Similarity: 100.000 Percent Identity: 100.000

FT: Alignment_block:

FT: US-09-080-127-2 x X00070 ..

FT: Align seg 1/1 to: X00070 from: 1 to: 1491

FT: 1 MetArgSerLeuLeuTrpAlaSerLeuLeuSerGlyValLeuAlaGlyAr 17

FT: 1 ATGAGTCCCTTTTGTGGCTTCGTTGCGGGCGTTTGGCTGGAG 50

FT: 17 GalLeuValSerProAspGluPheProGluAspIleGlnLeuGluAspL 34

|||||
51 GCGCGTTGTTGGCCGATGAGTTCGCCGAGATTTTACGTTGGAGATC 100
34 euLeuGluGlySerGlnGlnLeuGluAspPheAlaTyrAlaTyrProGlu 50
101 TCCTGGAAGATCCCAACAGCTTGAGACTTCGCTATGCTTACCCCGAG 150
51 ArgAsnArgValPheGlyGlyValAspAlaHisAspPheValAspTyrLe 67
151 CGCAATCGGCTTTGTTGTTGTTAAAGCCACAGCAGCAGGTTAACTATCT 200
67 uTyrGluGluLeuLysLysThrGlyTyrTyrAspValTyrLysGlnProg 84
201 CTACGAGGAGCGTGAAGAGACGTGCTACTATGATGCTACAAAGCAGCTC 250
84 InValHisLeuTrpSerAsnAlaAspGlnThrLeuLysValGlyAspGlu 100
251 AGGTCCACCTGTGGAGCAATGCCAGCAGCAGCTCAAGGTGGCGATGAG 300
101 GlnIleGluAlaLysThrMetThrTyrSerProSerValGluValThrAl 117
301 GAAATCGAGGCGAAGACCATGCTTACAGTCCACGCTCGAGGTCACCCG 350
117 aAspValAlaValLysAsnLeuGlyCysSerGluAlaAspTyrProS 134
351 CGATGTAGCCGTCGTCAGAACCTGGATGACGAGGCGGATTAACCAT 400
134 eAspValAlaGluLysValAlaLeuIleLysAlaGlyGluCysProPhe 150
401 CGGATGTGAGGCGAAGGTCCGCTGATCAAGCGGGAATGCCCGCTTC 450
151 GlyAspLysSerValLeuAlaAlaLysAlaLysAlaAlaAlaSerIleVal 167
451 GCGCGAAGTCCGCTTCGCTGCTGCCAAACCAAGCCGCGGCTGATTTGT 500
167 ITyrAsnAsnValAlaGlySerMetAlaGlyThrLeuGlyAlaAlaGln 184
501 CTATTAACATGTTGGCGGATCATGCGCGGCAACCTTGCGCGGCGAGA 550
184 eAspLysGlyProTyrSerAlaIleValAlaGlyIleSerLeuGluAsp 200
551 GGATTAAGGAGACCGATTCGCGCATTCGCTGATCAAGCTTGAGAGATGC 600
201 GlnLysLeuIleLysLeuAlaGluAlaGlySerValSerValAspLeuTr 217
601 CAGAGCTGATCAAGCTTCGAGGCTGATCGGATGTGGATCTGTG 650
217 pValAspSerLysGlnGluAsnArgThrThrTyrAsnValAlaGlnT 234
234 hTrpGlyGlyAspProAsnAsnValValAlaLeuGlyIleHisThrAsp 250
701 CGAAGGGGCGGATCCGAACACGTCGCGGCTGGGCGGACACGAGAC 750
251 SerValGluAlaGlyProGlyIleAsnAspAspGlySerGlyIleIleSe 267
751 TCAGTCGAGCGCGGCTGCTATCAACGACGATGGCTGGGCTATTATAG 800
267 AsnLeuValIleAlaLysAlaLeuThrGlnTyrSerValLysAsnAlav 284
801 CAACCTGTGATTCGCAAGCGCTCACGATACCTCCGTAAGAATGCGG 850
284 aLArgPheLeuPheThrPheAlaGluGluPheGlyLeuLeuGlySerAsn 300
851 TCGCTTCCTCTCTGACACAGAGGAGTTCGCTGCTGGACGAC 900
301 TyrTyrValSerHisLeuAsnAlaThrGluLeuAsnLysIleArgLeuTy 317
901 TACTACGCTCCCATCTGAATGCCAGCGAGCTGAACAAGATCCGACGTGA 950
317 rLeuAsnPheAspMetIleAlaSerProAsnTyrAlaLeuMetIleTyrA 334

```

951 CCTGACTTCGACATGATCGCCTCCTACTAGCCCTCATGATCATG 1000
334 spgIyAspGlySerAlaPheAsnGlnSerGlyProAlaGlySerAlaGln 350
|||||
1001 ACGGATGATGATCGGCGCTTCAACGACGAGCCGCGGCTTCCGCCG 1050
351 lIleGluYsLeuPheGluAspTyrThrAspSerIleAspIleuProHlaI 367
|||||
1051 ATCGAGAAAGCTTTCGAGGACTACTACGACTCCATCGACCTGCTCATAT 1100
367 eProThrGlnPheAspGlyIyAsrSerAspTyrGluAlaPheIleuAsnG 384
|||||
1101 CCCACCCAGTTGACGACGACCTTCGACTACGAGGCCCTTTATCCGAAC 1150
384 lYlIleProSerGlyLeuPheThrGlyAlaGluGlyIleMetSerGlu 400
|||||
1151 GCATTCCTCGCTGGTGACTCTTCAAGGCGCGGAGGCGCATGATGCCGA 1200
401 GluAsnAlaSerArgTyrGlyGlnAlaGlyAlaIaIaTyrAspAlaAs 417
|||||
1201 GAGAACCGCAACCCCGTGGGAGGTCAAGCCGCGGCTTACGAGCCGCA 1250
417 nTyRHsAlaAlaGlyAspAsnMetThrAsnLeuAsnHsGluAlaPheL 434
|||||
1251 CTACCAAGCCCGCGGAGACAAACATACCAACCTCAACCATGAGCTTCC 1300
434 euIleAsnSerLysAlaThrAlaPheAlaValaIaIaThrTyrAlaAsn 450
|||||
1301 TGATCAACTCCAAAGCCACCGCTTCGCGCGCCACCTACGCGCCACGAC 1350
451 LeuSerSerIleProLysArgAsnThrThrSerSerLeuHsArgArgAl 467
|||||
1351 CTCTCTCGATGCCCAAGGAAATACCAATCTCTTTCACCGACGAGAC 1400
467 aArgThrMetArgProPheGlyLysArgAlaProLysThrHsAlaHsI 484
|||||
1401 CCGACCATTCGACCATTCGGCAAGAGAGCTCCGAAAGACACCGCTCAG 1450
484 alSerGlySerGlyCysThrPheSerGlnValGluAla 496
|||||
1451 TATCAGGATCCGGATGCTGCAATCTCTCAAGTCGAGGA 1488
seq_name: N_Geneseq_36:V82521
seq_documentation_block:
ID V82521 standard; DNA; 1491 BP.
AC V82521;
DT 17-MAR-1999 (first entry)
DE Aspergillus oryzae aminopeptidase II encoding DNA.
KW aminopeptidase; protein hydrolysate; glycine releasing; protease;
KW protease; material; flavour; food; baking; animal feed additive;
KW palatability; hydrolysis; solubility; emulsifying; foaming; aroma;
KW mouthfeel; crust colour; ss.
OS Aspergillus oryzae.
FH Key location/Qualifiers
FT 1. 1491
FT CDS /*tag- a
PN WO9851163-A2.
PD 15-NOV-1998.
PF 15-MAY-1998; U09998.
PR 16-DEC-1997; US-069719.
PR 16-MAY-1997; US-857886.
PR 20-OCT-1997; US-062893.
PR 16-DEC-1997; DK-001465.
PA (ASAHI) ASAHI CHEM IND CO LTD.
PA (NOVO) NOVO NORDISK BIOTECH INC.
PA (NOVO) NOVO-NORDISK AS.
PI Blinovsky A, Brown K, Byun T, Fujii M, Golightly E,
PI Kofod Lv, Marumotac, Mathiasen TE;
PI WPI; 99-045177/04.
DR P-PSDB; W89586.
PT Production of protein hydrolysate - using protease and enzyme that
releases glycine, useful as flavour improvers in foods and animal

```

```

PT feed additives
PS Claim 4; Fig 1; 84pp; English.
CC A method has been developed for the production of protein hydrolysates
CC (PH) comprising reacting a protein with: (i) at least one polypeptide
CC with Gly-releasing activity; and (ii) at least one other protease so
CC that the amount of Gly produced is greater than when the protease is
CC used alone. PH are used to improve flavour of foods (e.g. baked goods)
CC and as animal feed additives. PH are preferably also enriched in Glu
CC (free and/or peptide bound), so have improved flavour and palatability.
CC Addition of a polypeptide with Gly-releasing activity increases the
CC degree of hydrolysis (or reduces the amount of enzyme needed) and
CC hydrolysates have better solubility, and emulsifying and foaming
CC properties. Baked goods containing them have improved aroma, mouthfeel
CC and crust colour. The present sequence encodes Aspergillus oryzae
CC aminopeptidase II, which is used in the method of the invention.
SQ Sequence 1491 BP; 335 A; 438 C; 426 G; 292 T;

alignment_scores:
Quality: 496.00 Length: 496
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-080-127-2 x V82521 ..

Align seg 1/1 to: V82521 from: 1 to: 1491

1 MetArgSerLeuLeuTyrPheAlaSerLeuLeuSerGlyValIleuAlaGlyArg 17
|||||
1 ATGAGTGCGCTTTGTGGGCTGCTGCTTTCGGCGCTGTGGCTGGGAG 50
17 gAlaLeuValSerProAspGluPheProGluAspIleGlnLeuGluAspL 34
|||||
51 GCGCCTGTTTCGCCGAGATGATGCCCGAGATATTCAGTTGGAATC 100
34 euIleuGluGlySerGlnGlnLeuGluAspPheAlaTyrAlaTyrProGlu 50
|||||
101 TGCTGGAGAGATCCCAACAGCTTACGACTTCGCTTACCTACCCGAG 150
101 tGcTGGAGAGATCCCAACAGCTTACGACTTCGCTTACCTACCCGAG 150
|||||
51 ArgAsnArgValPheGlyGlyLysAlaHsAspAspThrValaAsnTyrLe 67
|||||
151 CGCAATCGCGCTTGTGGTGAAGCCACGACGACACGAGTTACTACTATCT 200
67 uTyRGluGluLeuLysLysThrGlyTyrTyrAspValTyrLysGlnProG 84
|||||
201 CTACAGAGAGCTGAAGAGACTGCTACTATGATGTCTACAGCAGCCTC 250
84 lNValHsIleuTyrSerAsnAlaAspGlnThrLeuLysValAlaYAspGlu 100
|||||
251 AGGTGCACCTTGTGAGCAATGCCACAGACGCTCAAGAGTGGGAGATGAG 300
101 GIuIleGluAlaLysThrMetThrTyrSerProSerValGluValThrAl 117
|||||
301 GAAATCGAGCGGAGACCATGACTACAGTCCCGCTCGAGGCGACCGC 350
117 aAspValAlaValAlaLysAsnLeuGlyCysSerGlnAlaAspTyrProS 134
|||||
351 CGATGTAGCCGCTCGTCAAGAACCTGGGATGCGAGCGGATTAACCAT 400
134 eAspPValGluGlyLysValAlaLeuIleLysArgGlyLysCysProPhe 150
|||||
401 CCGATTCGAGGCGCAAGCTGCCCTGTATCAAGCTGGAAGATGCCCTTC 450
451 gGcGACAAAGTGGTTCCTCGCTGCCCAAGCCAGCGCCGCTTCGATTGT 500
151 GlYAspIySerValIleuAlaAlaLysAlaLysAlaAlaAspSerIleVa 167
|||||
451 GCGACAAAGTGGTTCCTCGCTGCCCAAGCCAGCGCCGCTTCGATTGT 500
167 lTyRAsnAsnValAlaGlySerMetAlaGlyThrLeuGlyAlaAlaGlnS 184
|||||
501 CTATTAACAATGTGGCCGATTCATGCGGCGACCTTGGCGGCGGCGAGA 550
184 eRAspIyGlyProTyrSerAlaIleValGlyIleSerLeuGluAspGly 200

```

```

|||||
551 GTATATAGGACCGTATTCGGCCATTCCTGGTATCAGCTTGAGAGATGCG 600
201 GlnLysLeuIleLysLeuAlaGluAlaGlySerValSerValAspLeuTr 217
|||||
601 CAGAGGTGATCAAGCTTGCTGAGGCGATCGGTATCTGATCTGCTGCTG 650
217 PValAspSerLysGlnGlnLysAsnArgTrpThrTrpAsnValAlaGln 234
|||||
651 GGTGGATAGTACAGAGAGAACCGTACAGAGTATACGTTGTCGCCGAGA 700
234 hTrpGlyGlyAspProAsnAsnValAlaLeuGlyGlyHisThrAsp 250
701 CGAAGGCGCGGATCCGAACAAGCTCTCCGCTGGGTGGCCACAGCGAC 750
251 SerValGluAlaGlyProGlyIleAsnAspAspGlySerGlyIleIle 267
|||||
751 TCAGTCGAGCGGGCCCTGGTATCAACGAGATGGCTCGGCAATTATTAG 800
267 rAsnLeuValIleAlaAlaLysAlaLeuThrGlnTrpSerValLysAsn 284
|||||
801 CAACCTGGTCTATTGCAAAAGCGCTCAGCGAGTACTCCGTCAAGAATGCC 850
284 AlaArgPheLeuPheTrpThrAlaGluGluPheGlyLeuLeuGlySer 300
|||||
851 TCGGCTTCCTCTCTGACAGCAGCAGAGAGTTCGCTCGCTGGCGACAGC 900
301 TyrTrpValSerHisLeuAsnAlaThrGluLeuAsnLysIleArgLeu 317
|||||
901 TACTAGCTCCCATCTGATGACCGACGAGTCAACAAGATCCGACTGTA 950
317 rLeuAsnPheAspMetIleLeuSerProAsnTrpAlaLeuMetIleTyr 334
|||||
951 CCTGAATTCGACATGATGCCCTCACCCTACCTACCTCATATATCATAG 1000
334 spLysAspGlySerAlaPheAsnGlnSerGlyProAlaGlySerAlaGln 350
1001 ACGGTATGATCGGCGTTCAACAGAGCGGCGCGGCTTCGGCCAG 1050
351 IleGluLysLeuPheGluAspTrpTrpAspSerIleAspLeuProHis 367
|||||
1051 ATCGAATAACTGTTGAGAGCTACTGACTGACTGACTGACTGCTCATAT 1100
367 eProThrGlnPheAspGlyArgSerAspTrpGluAlaPheIleLeuAsn 384
|||||
1101 CCCACCCGATTTGACGAGCTTCCACATACGAGCCCTTATCTTAAAG 1150
384 yLLeuProSerGlyGlyLeuPheThrGlyAlaGluGlyIleMetSerGlu 400
1151 GCATTCGCTCGGCTGAGCTCTTCAAGCGGCGCGAGGCAATCATGTCGA 1200
401 GlnAsnHisSerArgTrpGlyGlnAlaGlyValAlaTyrAspAlaAs 417
1201 GAGAAAGCGAAGCGCTGGGAGGTCAAGCGGCGGTACAGCCCA 1250
417 nTrpHisAlaAlaGlyAspAsnMetHisAsnLeuAsnHisGlyAlaPhe 434
1251 CTACCGCGCGCGGAGACACATGACCACTCAACCATGAAGCTTCC 1300
434 euLLeuSerLysAlaThrAlaPheAlaValAlaThrTyrAlaAsnAsp 450
|||||
1301 TATCAATCTCAAAAGCCACCGCTTGGCGTCCGACCTACGCGCAAGC 1350
451 LeuSerSerIleProLysArgAsnTrpTrpSerSerLeuHisArgArg 467
|||||
1351 CTTCTCTCATCTCCCAAGGATATACATCTCTCTTACACGAGAGAC 1400
467 ArgTrpMetArgProPheGlyLysArgAlaProLysThrHisAlaHis 484
1401 CCGCACCAAGCGACATTCGGCAAGAGAGTCCGAAGACACAGCTCAGC 1450
484 aLserGlySerGlyCysTrpHisSerGlnValGluAla 496
|||||

```

```

1451 TATCAGATCCGATCGCTGCAATCTCAAGTCGAGCA 1488
seq_name: N_Geneseq_36:X00018
seq_documentation_block:
ID X00018 standard: DNA; 1491 BP.
AC X00018:
DT 22-MAR-1999 (first entry)
DE Aspergillus oryzae aminopeptidase II encoding DNA.
KW Aminopeptidase; protein hydrolysate; baking; dough; flavour enhancer;
mouthfeel; palatability; aroma; hydrolysis; animal feed additive; ss.
OS Aspergillus oryzae.
FH Key
FT CDS
FT 1..1491 Location/Qualifiers
FT /*tag= a
FT /note= "at position 277 to 279 the DNA encodes gln
FT (position 93 in the protein) which corresponds
FT to the protein given in the sequence listing,
FT but the protein given in the figure has Asp at
FT this position"
FT
PD WO9851804-A1.
PD 19-NOV-1998.
PF 15-MAY-1998; U09940.
PR 20-OCT-1997; US-062893.
PR 16-MAY-1997; US-857886.
PA (NOVO ) NOVO NORDISK BIOTECH INC.
PI Blinkovsky A, Brown K, Byun T, Gollightly E, Kofod LV,
DR WPI: 99-045233/04.
DR P-PSDB: W89597.
PT New aminopeptidase from Aspergillus oryzae - used to produce protein
PT hydrolysates enriched in particular amino acids, used as flavour
PT enhancers, e.g. in doughs.
PS Claim 1, Fig 1; 92pp; English.
CC The present sequence encodes aminopeptidase (AP) II from Aspergillus
CC oryzae. AP is used in combination with an endopeptidase (EP) to
CC hydrolyse polypeptides, producing protein hydrolysate (PH), useful in
CC foods as flavour enhancer, e.g. in baked goods, enriched in: (a) Leu,
CC Gly (especially), Glu, Ser, Asp, Asn, Pro, Cys, Ala and/or Gln, or (b),
CC if the substrate has been deamidated, in Glu (free and/or peptide bound),
CC in which case products are useful as animal feed additives. AP can be
CC used in flavour-improving compositions (optionally containing EP) and in
CC dough pre-mixes. Also AP can be used for deactivating enzymes and for
CC converting precursors to mature proteins. AP increases the level of
CC hydrolysis of proteins and thus of flavour development. PH have improved
CC solubility, emulsifying and foaming properties, and products containing
CC them have better flavour, palatability and aroma.
CC Sequence 1491 BP; 335 A; 438 C; 426 G; 292 T;

```

alignment_scores:

Quality: 496.00 Length: 496
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-080-127-2 x X00018 ..

Align seg 1/1 to: X00018 from: 1 to: 1491

```

1 MetArgSerLeuLeuTrpAlaSerLeuLeuSerGlyValLeuAlaGlyArg 17
|||||
1 ARGAGGTGCGCTTTGTGGGCTTCGTGCTTGGCGCGGTGGCGGAG 50
17 GAlaLeuValSerProAspGluPheProGluAspIleGlnLeuGluAsp 34
|||||
51 GCGCGTTGTTGCGCGGATGAGTTCCCGAGAGATATTCAATTGGAAAGATC 100
34 euLeuGlnGlySerGlnGlnLeuGlnAspPheAlaTyrAlaTyrProLys 50
|||||
101 TCTTGAAAGATCCCAACAGCTTGAGACTTGCGCTTACGCTTACCCGAG 150
51 ArgAsnArgValPheGlyGlyLysAlaHisAspThrValaAsnTyrLe 67
|||||

```

```

151 CGCAATGCGCTTTGGTGGTAAAGCCACGACGACGATTACTATCT 200
67 uTYRGLuLeuLysLysThrgLYTYRAspValTYRlySGlnProg 84
|||||
201 CTACGAGAGACTGAAAGAGACTGGCTACTATGATGCTACAGACGACCTC 250
84 lNvalHsLeuTrpSerAsnAlaAspGlnHrLeuLysValGlyAspGlu 100
|||||
251 AGGIGACCTGTGGAGCAATGCCACGACGAGCTCAAGGTGGGAGATGAG 300
101 GluileGluAlaLysThrmethTrpYrSerProSerValGluValHrAl 117
|||||
301 GAATATGAGGCGAAGACCATGACTACAGTCCCGCTCGAGGTACACGCC 350
117 aAspValAlaValAlaLysAsnLeuGlyCysSerGluAlaAspTYrPro 134
|||||
351 CGATGTAGCCGCTGCAAGAACCTGGGATGCAGCGAGCGGATTAACCAT 400
134 eRAspValGluGlyValAlaLeuHrLeuLysArgGlyGlyCysProPhe 150
|||||
401 CCGATGTGAGGCGAAGGTGCGCTGATCAAGCGTGAAGATGCCGCTTC 450
151 GlyAspLysSerValLeuAlaAlaLysAlaLysAlaAlaSerIleVal 167
|||||
451 GCGGACAGAGTCGGTTCGCTGCCAAAGCCAGGCGCGGCTTCGATTGT 500
167 lTYrAsnAsnValAlaGlySerMetAlaGlyThrLeuGlyAlaAlaGln 184
501 CTATTAACAATGTGGCGGATCCATGCGGCGCACCCCTTGGCGCGCGAGA 550
184 eRAspLysGlyProTYrSerAlaIleValGlyIleSerLeuGluAspGly 200
|||||
551 GTGATTAAGGACCGTATTCGCGCATTCGATTCAGCTTGGAGAGATGAGC 600
201 GlnLysLeuLysLeuAlaGluAlaGlySerValSerValAspLeuTr 217
|||||
601 CAGAGGTGATCAAGCTTGCTGAGCGCTGATCTGTGATCTGTGTG 650
217 pValAspSerLysGlnGluAsnArgThrTYrAsnValAlaAlaGlnT 234
|||||
651 GGTGGATAGTAAACAGAGAACCTGACGAGTAAACGTTGTCCCGCAGA 700
234 hrLYSGlyGlyAspProAsnAsnValAlaLeuGlyGlyHsThrAsp 250
|||||
701 CGAAGGCGCGGATCCGACACACCTCGTGGGCTGGGTGGCCACACGAC 750
251 SerValGluAlaGlyProGlyIleAsnAspAspGlySerGlyIleIle 267
|||||
751 TCAGTGCAGGCGGGCCCTGGTATCAACGACGATGCTCGGGCATTTAG 800
267 rAsnLeuValIleAlaLysAlaLeuThrGlnTYrSerValLysAsnAla 284
|||||
801 CAACCTGGTATTGCCAAAGCGCTCACGCACTACTCCGTCAAGATGCCG 850
284 aLArgPheLeuPheTrpThrAlaGluLupheGlyLeuLysIleAsn 300
|||||
851 TGGCTTCCTCTTGTGGACAGCAGAGAGTGGCTGTGCTGGGAGCAAC 900
301 TYrTYrValSerHsLeuAsnAlaThrGluLeuAsnLysIleArgLeuTY 317
|||||
901 TACTAGCTGCCCATCTGATGCCACCGAGCTGAACAAGATCGACATGTA 950
317 rLeuAsnPheAspMetIleAlaSerProAsnTYrAlaLeuMetIleTYr 334
|||||
951 CCGAATTTGACATGATCCCTCACCTACTACGCCCTCATGATCTATG 1000
334 spGlyAspGlySerAlaPheAsnGlnSerGlyProAlaGlySerAlaGln 350
|||||
1001 ACGGTGATGATCGGCTTCAACCAAGCGGACCGGCTTCGCCCCAG 1050
351 lIleGluLysLeuPheGluAspTYrTYrAspSerIleAspLeuProHsI 367
|||||
1051 ATCGAGAAACTGTTCGAGACTACTACGACTCCATCGACTGCTCATAT 1100

```

```

367 eProThrGlnPheAspGlyArgSerAspTYrGluAlaPheIleLeuAsn 384
|||||
1101 CCCACCCAGTTTGCAGGAGCTTCGACTACGAGGCGCTTATCTTAACG 1150
384 lYIleProSerGlyGlyLeuPheThrGlyValAlaGlyIleMetSerGlu 400
|||||
1151 GCATTCGCTCGGTGAGCTTTCAAGGCGCCGAGGGCATCATGTCCGA 1200
401 GluAsnAlaSerArgTrpGlyGlnAlaGlyValAlaAlaTYrAspAla 417
|||||
1201 GAGAGCCAGCAAGCCGTGGGAGGTCAAGCGGCTGGCTACAGCCCA 1250
417 nTYrHsAlaAlaGlyAspAsnMetHrAsnLeuAsnHsGluAlaPhe 434
|||||
1251 CTACACAGCGCGCGGAGACAACTGACCAACCTCAACATGAAGCCCTCC 1300
434 euIleAsnSerLysAlaThrAlaPheAlaValAlaThrTYrAlaAsnAsp 450
|||||
1301 TGATCACTCCAAAGCCACCGCTTGCCTGCGCACCTACGACCAAGAC 1350
451 LeuSerSerIleProLysArgAsnThrTrpSerSerLeuHsArgArgAl 467
|||||
1351 CTCCTCCGATCCCAAGCAAGGATACACACTCTCTTGACCGAGAGAGC 1400
467 aArgThrMetArgProPheGlyLysArgAlaProLysThrHsAlaHsY 484
|||||
1401 CCCGACCATCGACATTCGCGCAAGAGAGCTCCGAAGACACAGCTCAG 1450
484 aLserGlySerGlyCysTrpHsSerGlnValGluAla 496
|||||
1451 TATCAGGATCCGATGCTGCAATTCAGTCAAGTCAGGCA 1488

seq_name: N_Geneseq_36.X00020

seq_documentation_block:
ID X00020 standard; DNA; 36 bp.
AC X00020:
DT 17-MAR-1999 (first entry)
DE Aspergillus oryzae aminopeptidase II PCR reverse primer.
KW Aminopeptidase; protein hydrolysis; baking; dough; flavour enhancer;
KW mouthfeel; palatability; aroma; hydrolysis; animal feed additive;
KW PCR primer; ss.
OS Synthetic.
OS Aspergillus oryzae.
FH Key Location/Qualifiers
FT modified_base 7
FT /*tag= a
/mod_base= 1
/note= "inosine"
FT modified_base 10
FT /*tag= b
/mod_base= 1
/note= "inosine"
FT modified_base 13
FT /*tag= c
/mod_base= 1
/note= "inosine"
FT modified_base 16
FT /*tag= d
/mod_base= 1
/note= "inosine"
FT modified_base 22
FT /*tag= e
/mod_base= 1
/note= "inosine"
FT modified_base 25
FT /*tag= f
/mod_base= 1
/note= "inosine"
FT modified_base 27
FT /*tag= g
/mod_base= 1

```

```

FT modified_base 34 /note- "inosine"
FT FT /*tag- h
FT /mod_base- 1
FT /note- "inosine"
PN W09851804-A1.
PD 19-NOV-1998.
PF 15-MAY-1998; U09940.
PR 20-OCT-1997; US-062893.
PR 16-MAY-1997; US-857886.
PA (NOVO ) NOVO NORDISK BIOTECH INC.
PI Blinovsky A, Brown K, Byun T, Golightly E, Kofoed LV;
DR WPI: 99-045233/04.
PT New aminopeptidase from Aspergillus oryzae - used to produce protein
PT hydrolysates enriched in particular amino acids, used as flavour
PT enhancers, e.g. in doughs
PS Example 6; Page 49; 92pp; English.
CC The present sequence represents a PCR primer for aminopeptidase
CC (AP) II. AP is used in combination with an endopeptidase (EP) to
CC hydrolyse polypeptides, producing protein hydrolysate (PH), useful in
CC foods as flavour enhancer, e.g. in baked goods, enriched in: (a) Leu,
CC Gly (especially), Glu, Ser, Asp, Asn, Pro, Cys, Ala and/or Gln, or (b),
CC if the substrate has been deamidated, in Glu (free and/or peptide bound),
CC in which case products are useful as animal feed additives. AP can be
CC used in flavour-improving compositions (optionally containing EP) and in
CC dough pre-mixes. Also AP can be used for deactivating enzymes and for
CC converting precursors to mature proteins. AP increases the level of
CC hydrolysis of proteins and thus of flavour development. PH have improved
CC solubility, emulsifying and foaming properties, and products containing
CC them have better flavour, palatability and aroma.
SQ Sequence 36 BP; 5 A; 9 C; 3 G; 7 T;

alignment_scores:
Quality: 12.00 Length: 12
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-080-127-2 x X00020/rev ..

Align seg 1/1 to reverse of: X00020 from: 1 to: 36

113 ValGIuValIhAlaSPValAlaValValLysAsn 124
|||||
36 GTNGARGTNACNGCNGAYGTNGCNGTNGTNAARAAY 1

seq_name: N_Geneseq_36:V82524
seq_documentation_block:
ID V82524 standard: DNA; 36 BP.
AC V82524;
DT 17-MAR-1999 (first entry)
DE Aspergillus oryzae aminopeptidase II PCR reverse primer.
KW Aminopeptidase; protein hydrolysate; glycine releasing; protease;
KW proteinaceous material; flavour; food; baking; animal feed additive;
KW palatability; hydrolysis; solubility; emulsifying; foaming; aroma;
KW mouthfeel; crust colour; PCR primer; ss.
OS Synthetic.
OS Aspergillus oryzae.
FH key Location/Qualifiers
FT modified_base 7
FT /*tag- a
FT /mod_base- 1
FT /note- "inosine"
FT 10
FT /*tag- b
FT /mod_base- 1
FT /note- "inosine"
FT 13
FT /*tag- c
FT /mod_base- 1
FT /note- "inosine"
..FT

```

```

FT modified_base 16
FT FT /*tag- d
FT /mod_base- 1
FT /note- "inosine"
FT 22
FT modified_base 22
FT FT /*tag- e
FT /mod_base- 1
FT /note- "inosine"
FT 25
FT modified_base 25
FT FT /*tag- f
FT /mod_base- 1
FT /note- "inosine"
FT 28
FT modified_base 28
FT FT /*tag- g
FT /mod_base- 1
FT /note- "inosine"
FT 34
FT modified_base 34
FT FT /*tag- h
FT /mod_base- 1
FT /note- "inosine"
PN W09851163-A2.
PD 19-NOV-1998.
PF 15-MAY-1998; U09998.
PR 16-DEC-1997; US-069719.
PR 16-MAY-1997; US-857886.
PR 20-OCT-1997; US-062893.
PR 16-DEC-1997; DK-001465.
PA (ASAHI ) ASAHI CHEM IND CO LTD.
PA (NOVO ) NOVO NORDISK BIOTECH INC.
PI Blinovsky A, Brown K, Byun T, Fujii M, Golightly E,
PI Kofoed LV, Marumotac, Mathiansen TE;
DR WPI: 99-045177/04.
PT Production of protein hydrolysate - using protease and enzyme that
PT releases glycine, useful as flavour improvers in foods and animal
PT feed additives
PS Example 6; Page 23; 84pp; English.
CC A method has been developed for the production of protein hydrolysates
CC (PH) comprising reacting a protein with: (i) at least one polypeptide
CC with gly-releasing activity; and (ii) at least one other protease so
CC that the amount of Gly produced is greater than when the protease is
CC used alone. PH are used to improve flavour of foods (e.g. baked goods)
CC as animal feed additives. PH are preferably also enriched in Glu
CC (free and/or peptide bound), so have improved flavour and palatability.
CC Addition of a polypeptide with Gly-releasing activity increases the
CC degree of hydrolysis (or reduces the amount of enzyme needed) and
CC hydrolysates have better solubility, and emulsifying and foaming
CC properties. Baked goods containing them have improved aroma, mouthfeel
CC and crust colour. The present sequence represents a PCR primer for
CC Aspergillus oryzae aminopeptidase II, which is used in the method of
CC the invention.
SQ Sequence 36 BP; 5 A; 9 C; 3 G; 7 T;

alignment_scores:
Quality: 12.00 Length: 12
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-080-127-2 x V82524/rev ..

Align seg 1/1 to reverse of: V82524 from: 1 to: 36

113 ValGIuValIhAlaSPValAlaValValLysAsn 124
|||||
36 GTNGARGTNACNGCNGAYGTNGCNGTNGTNAARAAY 1

seq_name: N_Geneseq_36:T90735
seq_documentation_block:
ID T90735 standard: DNA; 393 BP.
AC T90735;

```

DE 20-APR-1998 (first entry)
 DE EST_R35464.DNA.
 KW Human; consensus bikunin; inhibition; trypsin; kallikrein;
 KW plasmin; factor XIIa; treatment; prevention; oedema;
 KW inflammation; infection; granulomatosis; multiple sclerosis;
 KW ischaemia; perioperative blood loss; sepsis; shock; fibrosis;
 KW blood coagulation disease; polytrauma; stroke; haemorrhage;
 KW gastric cancer; cervical cancer; metastasis; blood loss;
 KW EST_R35464; ss.
 OS Homo sapiens.
 FT Key Location/Qualifiers
 FT CDS 3..392
 FT /tag= a
 FT WO9733996-A2.
 PD 18-SEP-1997.
 PF 10-MAR-1997; U03894.
 PR 04-OCT-1996; US-725251.
 PR 11-MAR-1996; US-013106.
 PR 14-JUN-1996; US-019793.
 PA (FARB) BAYER CORP.
 PI Davis G, Delaria KA, Marlor CW, Muller DK, Tamburini PP;
 DR WPI: 97-470876/43.
 DR P-PSDB; W30062.
 PT New human placental bikunin - used to inhibit kallikrein, trypsin
 PT etc. in treatment of oedema, multiple sclerosis, fibrosis, or
 PT perioperative blood loss
 PS Disclosure: Fig 1; 110pp; English.
 CC The present sequence is the EST_R35464 DNA, which is similar to
 CC human bikunin. Bikunin inhibits, e.g. trypsin, kallikrein, plasmin
 CC and factor XIIIa.
 CC Bikunin can be used to treat or prevent brain and spinal cord
 CC oedema, inflammation, infection or granulomatosis, multiple
 CC sclerosis, ischaemia, perioperative blood loss, sepsis, shock,
 CC fibrosis, blood coagulation diseases, polytrauma, stroke,
 CC cerebral or subarachnoid haemorrhage and gastric or cervical
 CC cancer and prevent metastasis. It is particularly useful for
 CC reducing blood loss during surgery, and can also be used to treat
 CC other cancer, arthritis, anaemia, non-insulin dependent diabetes,
 CC influenza and similar viral infections, acute pancreatitis and
 CC gout, and prevent pre-term labour. It has similar properties to
 CC apolitin, but is less highly charged so should be less
 CC immunogenic and less likely to damage the kidneys. Manipulation
 CC of the bikunin sequence may allow the inhibitory profile to be
 CC altered. It also reduces or eliminates the need for whole donor
 CC blood or blood products during surgery, thereby reducing the risk
 CC of infection and other adverse side effects, as well as reducing
 CC the cost of surgery.
 SO Sequence 393 BP; 85 A; 103 C; 115 G; 86 T;

alignment_scores:
 Quality: 9.00 Length: 9
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-080-127-2 x T90735 ..

Align seg 1/1 to: T90735 from: 1 to: 393

9 LeuLeuSerGlyValLeuAlaIArg 17
 ||||||||||||||||||||
 35 CTCCTCTCTGGGGCTGTGGCCGGCCGA 61

seq_name: N_Geneseq_36:065520

seq_documentation_block:
 ID 065520 standard; CDNA: 2653 BP.
 AC 065520;
 DT 11-JAN-1995 (first entry) antigen cDNA.
 DE Prostate-specific membrane antigen PSM; prostate cancer;
 KW Prostate-specific membrane antigen; PSM; targeting; tumour detection;
 KW transmembrane glycoprotein; imaging; targeting; tumour detection;

KW antibody detection; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 262..2514
 FT /tag= a
 FT /product= prostate specific membrane antigen (PSM)
 PN WO9409820-A.
 PD 11-MAY-1994.
 PF 05-NOV-1993; U10624.
 PR 05-NOV-1992; US-973337.
 PA (SLOK) SLOAN KETTERING INST CANCER.
 PI Fair WR, Heston MDW, Israeli RS;
 DR WPI: 94-167129/20.
 DR P-PSDB; R55097.
 PT Prostate-specific membrane antigen and DNA encoding it - is
 PT useful for detecting haematogenous micro-metastatic tumour cells
 PT and for identifying ligands which bind to PSM Ag
 PS Claim 3: Page 103-106; 196pp; English
 CC 065520 encodes a prostate specific membrane antigen (PSM, P55097).
 CC The PSM coding sequence is useful for suppressing or modulating the
 CC metastatic ability of prostate tumour cells to grow, or for
 CC eliminating them. The protein is useful to identify or purify ligands
 CC of the Ag. It is also an attractive target for Ab-directed imaging
 CC and targeting of prostatic tumour deposits.
 SO Sequence 2653 BP; 782 A; 524 C; 640 G; 707 T;

alignment_scores:
 Quality: 9.00 Length: 9
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-080-127-2 x Q65520 ..

Align seg 1/1 to: Q65520 from: 1 to: 2653

291 AlaGluGluPheGlyLeuGlySer 299
 ||||||||||||||||||||
 1528 GCAGAGAAAGATTGTCTCTGTCTT 1554

seq_name: N_Geneseq_36:T36785

seq_documentation_block:
 ID T36785 standard; CDNA: 2654 BP.
 AC T36785;
 DT 04-NOV-1996 (first entry)
 DE Prostate-specific membrane antigen cDNA.
 KW Prostate-specific membrane antigen; PSM; promoter; prostate cancer;
 KW metastasis; gene therapy; diagnosis; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FH 5'utr 1..261
 FT /tag= a
 FT 262..2253
 FT /tag= b
 FT /product= PSM antigen
 FT misc_feature 114..380
 FT /tag= c
 FT /note= "bases 114-380 (-147 to +109) are absent
 FT in PSM' cDNA"
 FT polyA_signal 2352..2357
 FT /tag= d
 FT WO9626272-A1.
 PD 29-AUG-1996.
 PF 23-FEB-1996; U02424.
 PR 24-FEB-1995; US-394152.
 PR 02-JUN-1995; US-470735.
 PR 02-JUN-1995; US-466381.
 PA (SLOK) SLOAN KETTERING INST CANCER RES.
 PI Fair WR, Heston MDW, Israeli RS;
 DR WPI: 96-402365/40.
 DR P-PSDB; W02234.

PT DNA encoding alternatively spliced prostate-specific membrane
PT antigen - useful to develop prods. for detecting haematogenous
PT micrometastatic tumour cells, or prostate cancer progression
PS Example 1: Fig 47A-D: 284bp; English.
CC A CDNA clone (T36785) codes for human 100 kDa prostate-specific
CC membrane (PSM) antigen (M02234), an integral membrane glycoprotein
CC that is very highly expressed in prostatic tumours and metastases.
CC It was obt. from lymph node carcinoma of prostate (LNCaP) cell
CC mRNA by PCR amplification (see also T36795-808) and screening of an
CC LNCaP cDNA library using an amplified cDNA partial clone as probe.
CC The CDNA can be used to provide probes and primers useful e.g. in
CC detecting haematogenous micrometastatic tumour cells and determining
CC prostate cancer progression (see also T36827-30 and T36813-18).
CC and in gene therapy. An alternatively spliced PSM, PSM', has a
CC shorter CDNA sequence. PSM genomic DNA is given in T36786.
SQ Sequence 2654 BP; 782 A; 525 C; 639 G; 708 T;

alignment_scores:
Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-080-127-2 x T36785 ..

Align seg 1/1 to: T36785 from: 1 to: 2654

291 AlaglugluPhleglyLeuGlySer 299
|||||
1528 GCAGAGAAATGTCCTCTGCTGCTCT 1554

seq_name: N_Geneseq_36:T36785

seq_documentation_block:
ID V75088 standard; DNA: 1145 BP.
AC V75088;
DT 16-MAR-1999 (first entry)
DE Staphylococcus aureus confg SEQ ID #777.
KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.
OS Staphylococcus aureus.
FH Key Location/Qualifiers
FT misc-feature 481..540
FT /tag- a
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"

PN EP-786519-A2.
PD 30-JUL-1997.
PE 07-JAN-1997; 100117.
PR 05-JAN-1996; US-009861.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,
PI Rosen CA.
PI WPI; 97-374922/35.
DR Polynucleotide(s) and proteins derived from Staphylococcus aureus
PT stored on computer readable medium and used in the production of
PT anti-S.aureus vaccines
PS Claim 1: Page 1661-1662: 3271pp; English.
CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of

CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the
CC computer readable medium.
SQ Sequence 1145 BP; 342 A; 167 C; 145 G; 430 T;

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-080-127-2 x V75088 ..

Align seg 1/1 to: V75088 from: 1 to: 1145

238 AsProAsnAsnValAlaLeu 245
|||||
541 GATCCAAATATGTCCTTGA 564

seq_name: N_Geneseq_36:X34119

seq_documentation_block:
ID X34119 standard; DNA: 1503 BP.
AC X34119;
DT 06-JUL-1999 (first entry)
DE Mycobacterium species nucleic acid sequence 19D.
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW hybridisation; detection; vaccine; immunisation; infection; ss.
OS Mycobacterium sp.
PN MO9909186-A2.
PE 25-FEB-1999.
PF 14-AUG-1998; F01813.
PR 11-SEP-1997; FR-011325.
PR 14-AUG-1997; FR-010404.
PA (INSP) INST PASTEUR.
PI Gicquel B, Lhm EM, Pellicic V, Portnoi D, Goguet de la Salmoniere Y,
PI Guigueno A;
PI WPI; 99-181045/15.
DR P-P-SDB; Y04867.
DE Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in
PT infection-associated protein expression
PS Claim 22; Fig 19D; 309pp; French.
CC Sequences X34001-X34252 represent nucleic acids encoding secreted
CC proteins from various Mycobacterium species microorganisms. The
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
SQ Sequence 1503 BP; 275 A; 470 C; 502 G; 256 T;

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-080-127-2 x X34119 ..

Align seg 1/1 to: X34119 from: 1 to: 1503

384 GlyTleProSerglyGlyLeu 391
|||||
1237 GGTATCCCTTCGGGTGCTGTC 1260

seq_name: N_Geneseq_36:X34120

seq_documentation_block:
 ID X34120 standard; DNA; 1587 BP.
 AC X34120;
 DT 06-JUL-1999 (first entry)
 DE Mycobacterium species nucleic acid sequence 19F.
 KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
 hybridisation; detection; vaccine; immunisation; infection; ss.
 OS Mycobacterium sp.
 PN WO9909186-A2.
 PD 25-FEB-1999.
 PF 14-AUG-1998; F01813.
 PR 11-SEP-1997; FR-011325.
 PR 14-AUG-1997; FR-010404.
 PA (INSP) INST PASTER.
 PI Gicquel B, Lim EM, Pelletier V, Portnoi D, Gouget de la Salmoniere Y,
 DR P-PSDB: Y04868.
 DT 99-181045/15.
 PT Mycobacterial DNA vectors containing reporter constructs - for
 identifying coding or promoter sequences involved in
 infection-associated protein expression
 PS Claim 22; Fig 19F; 309pp; French.
 CC Sequences X34001-X34252 represent nucleic acids encoding secreted
 CC proteins from various Mycobacterium species microorganisms. The
 CC nucleotide sequences can be used as primers and probes for methods
 CC for detecting and identifying mycobacteria, especially belonging to
 CC the M. tuberculosis complex. The encoded proteins can be used in
 CC vaccines for immunisation against a bacterial or viral infection.
 SQ Sequence 1587 BP; 288 A; 490 C; 533 G; 276 T;
 seq_name: N_Geneseq_36:087422

alignment_scores:
 Quality: 8.00 Length: 8
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-080-127-2 x X34120 ..
 Align seg 1/1 to: X34120 from: 1 to: 1587

384 GLYIleProSerGLYleuPhe 391
 ||||||||||||||||||||
 1321 GGAATCCCTTCGGTGCCCTGTTTC 1344
 seq_name: N_Geneseq_36:087422

seq_documentation_block:
 ID 087422 standard; DNA; 32 BP.
 AC 087422;
 DT 16-NOV-1995 (first entry)
 DE Human GRK sense primer.
 KW Polymerase chain reaction; PCR; amplify; primer; isolation; beta-ARK1;
 KW G protein-coupled receptor kinase; GRK6; catalytic domain; detection;
 KW rhodopsin kinase; RK; beta adrenergic receptor kinase; immunisation;
 KW conserved amino acid; subdomain II; subdomain VII; quantification;
 KW signal transduction; immune system; inflammation; human; rat; ss.
 OS Synthetic.
 PN WO9507990-A.
 PD 23-MAR-1995.
 PF 16-SEP-1994; U10487.
 PR 17-SEP-1993; US-123932.
 PR 31-MAR-1994; US-221817.
 PA (ICOS-) ICOS CORP.
 PI Chantry D, Gray PW, Hoekstra MF;
 DR WPI: 95-131358/17.
 PT New G protein-coupled receptor kinase, GRK6, and corresp. DNA - are
 PT used in the prodn. of GRK6 and for the immunisation, treatment and
 PT diagnosis of diseases of the immune system and inflammation.
 PS Example 1; Page 9; 71pp; English.
 CC The sequences given in 087422-23 are primers which were used in the
 CC isolation of the novel human G protein-coupled receptor kinase (GRK6).

CC These primers are based on a partial rat GRK6 sequence isolated from
 CC rat T cell cDNA. These primers were used to amplify cDNA isolated
 CC from a human macrophage cDNA library. The isolated cDNA sequence was
 CC originally thought to have the sequence given in 087425 but on further
 CC isolation the sequence was found to have the sequence given in 087426.
 CC This sequence was found to have one amino acid difference, corresponding
 CC to residue 61, compared to the GRK6 sequence isolated from a human heart
 CC cDNA library by Benovic et al. J. Biol. Chem. (1993). Peptide fragments
 CC of GRK6 may be used in immunisation, detection and quantification of
 CC GRK6. They may also be used for modulating the activities of GRK6, esp.
 CC in treatment of diseases and physiological conditions of the immune
 CC system, such as inflammation.
 SQ Sequence 32 BP; 6 A; 7 C; 12 G; 7 T;
 seq_name: N_Geneseq_36:087422

alignment_scores:
 Quality: 7.00 Length: 7
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-080-127-2 x 087422 ..
 Align seg 1/1 to: 087422 from: 1 to: 32

117 AlaAspValAlaValLys 123
 ||||||||||||||||||
 12 GCGGACGTGGCGCGTCAAG 32
 seq_name: N_Geneseq_36:087423

seq_documentation_block:
 ID 087423 standard; DNA; 32 BP.
 AC 087423;
 DT 16-NOV-1995 (first entry)
 DE Human GRK antisense primer.
 KW Polymerase chain reaction; PCR; amplify; primer; isolation; beta-ARK1;
 KW G protein-coupled receptor kinase; GRK6; catalytic domain; detection;
 KW rhodopsin kinase; RK; beta adrenergic receptor kinase; immunisation;
 KW conserved amino acid; subdomain II; subdomain VII; quantification;
 KW signal transduction; immune system; inflammation; human; rat; ss.
 OS Synthetic.
 PN WO9507990-A.
 PD 23-MAR-1995.
 PF 16-SEP-1994; U10487.
 PR 17-SEP-1993; US-123932.
 PR 31-MAR-1994; US-221817.
 PA (ICOS-) ICOS CORP.
 PI Chantry D, Gray PW, Hoekstra MF;
 DR WPI: 95-131358/17.
 PT New G protein-coupled receptor kinase, GRK6, and corresp. DNA - are
 PT used in the prodn. of GRK6 and for the immunisation, treatment and
 PT diagnosis of diseases of the immune system and inflammation.
 PS Example 1; Page 9; 71pp; English.
 CC The sequences given in 087422-23 are primers which were used in the
 CC isolation of the novel human G protein-coupled receptor kinase (GRK6).
 CC These primers are based on a partial rat GRK6 sequence isolated from
 CC rat T cell cDNA. These primers were used to amplify cDNA isolated
 CC from a human macrophage cDNA library. The isolated cDNA sequence was
 CC originally thought to have the sequence given in 087425 but on further
 CC isolation the sequence was found to have the sequence given in 087426.
 CC This sequence was found to have one amino acid difference, corresponding
 CC to residue 61, compared to the GRK6 sequence isolated from a human heart
 CC cDNA library by Benovic et al. J. Biol. Chem. (1993). Peptide fragments
 CC of GRK6 may be used in immunisation, detection and quantification of
 CC GRK6. They may also be used for modulating the activities of GRK6, esp.
 CC in treatment of diseases and physiological conditions of the immune
 CC system, such as inflammation.
 SQ Sequence 32 BP; 7 A; 6 C; 11 G; 8 T;

alignment_scores:

Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-080-127-2 x Q87423 ..

Align seg 1/1 to: Q87423 from: 1 to: 32

117 Alaaspvalalavalvalys 123

|||||
12 GCGGACGTGGCGTGTCAAG 32

seq_name: N_Geneseq_36:Q87424

seq_documentation_block:

ID Q87424 standard; DNA; 32 BP.
AC Q87424;
DT 16-NOV-1995 (first entry)
DE Rat GRK DNA fragment.
KW Polymerase chain reaction; PCR; amplify; primer; isolation; beta-ARK1;
KW G protein-coupled receptor kinase; GRK6; catalytic domain; detection;
KW rhodopsin kinase; RK; beta adrenergic receptor kinase; immunisation;
KW conserved amino acid; subdomain II; subdomain VII; quantification;
KW signal transduction; immune system; inflammation; human; rat; ss.
OS Rattus rattus.
PN W055079907A.
PD 23-MAR-1995.
PE 16-SEP-1994; U10487.
PR 17-SEP-1993; US-123932.
PR 31-MAR-1994; US-221817.
PA (ICOS-) ICOS CORP.
PI Chantry D. Gray PW. Hoekstra MF.
PI WPI: 95-131358/17.
PT New G protein-coupled receptor kinase, GRK6, and corresp. DNA - are
PT used in the prodn. of GRK6 and for the immunisation, treatment and
PT diagnosis of diseases of the immune system and inflammation.
PS Example 1; Page 23; 71pp; English.
CC This sequence represents a fragment of the rat G protein-coupled
CC receptor kinase (GRK6) gene. This sequence was used in the design of
CC the sequences given in Q87422-23 which are primers which were used in
CC the isolation of the novel human G protein-coupled receptor kinase
CC (GRK6) gene. These primers were used to amplify cDNA isolated from a
CC human macrophage cDNA library. The isolated cDNA sequence was originally
CC thought to have the sequence given in Q87425 but on further isolation,
CC the sequence was found to have the sequence given in Q87426. This
CC sequence was found to have one amino acid difference, corresponding
CC to residue 61, compared to the GRK6 sequence isolated from a human heart
CC cDNA library by Benovic et al. J. Biol. Chem. (1993). Peptide fragments
CC of GRK6 may be used in immunisation, detection and quantification of
CC GRK6. They may also be used for modulating the activities of GRK6, esp.
CC those involved in signal transduction. Modulators of GRK6 may be useful
CC in treatment of diseases and physiological conditions of the immune
CC system, such as inflammation.
SQ Sequence 32 BP; 6 A; 7 C; 12 G; 7 T;

alignment_scores:

Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-080-127-2 x Q87424 ..

Align seg 1/1 to: Q87424 from: 1 to: 32

117 Alaaspvalalavalvalys 123

|||||
12 GCGGACGTGGCGTGTCAAG 32

seq_name: N_Geneseq_36:T24875

seq_documentation_block:

ID T24875 standard; cDNA to mRNA; 182 BP.
AC T24875;
DT 30-OCT-1996 (first entry)
DE Human gene signature HUMGS06981.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PE 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PI (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1716; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 182 BP; 35 A; 48 C; 59 G; 37 T;

alignment_scores:

Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-080-127-2 x T24875 ..

Align seg 1/1 to: T24875 from: 1 to: 182

242 valvalalaleuglyglyhis 248

|||||
114 GTGTGGCCCTTNGGGGCCAT 134

Thu May 11 10:03:04 2000

us-09-080-127-2.rng

Page 17

[illegible]

[illegible]

/cgn2_6/p/odata/2/1na/5B_CONB.seq:US-08-726-725-4 +	6.00	86.53	2.3e+03	759
/cgn2_6/p/odata/2/1na/5D_CONB.seq:US-08-833-659A-18 +	6.00	86.53	2.3e+03	759
/cgn2_6/p/odata/2/1na/5A_CONB.seq:US-08-276-552A-49 +	6.00	86.52	2.3e+03	760
/cgn2_6/p/odata/2/1na/5C_CONB.seq:US-08-798-744-49 +	6.00	86.52	2.3e+03	760
/cgn2_6/p/odata/2/1na/5A_CONB.seq:US-07-941-523-5 +	6.00	86.41	2.4e+03	771
/cgn2_6/p/odata/2/1na/5A_CONB.seq:US-07-941-523-5 +	6.00	86.38	2.4e+03	774
/cgn2_6/p/odata/2/1na/5D_CONB.seq:US-08-961-523-7 +	6.00	86.38	2.4e+03	774
/cgn2_6/p/odata/2/1na/5D_CONB.seq:US-08-762-508-9 +	6.00	86.38	2.4e+03	774
/cgn2_6/p/odata/2/1na/5B_CONB.seq:US-09-358-171-19 -	6.00	86.36	2.4e+03	776
/cgn2_6/p/odata/2/1na/6_CONB.seq:US-09-090-947-11 -	6.00	86.36	2.4e+03	776
/cgn2_6/p/odata/2/1na/5A_CONB.seq:US-07-941-523-3 +	6.00	86.35	2.4e+03	777
/cgn2_6/p/odata/2/1na/5D_CONB.seq:US-08-929-302-3 +	6.00	86.13	2.4e+03	800
/cgn2_6/p/odata/2/1na/5D_CONB.seq:US-08-929-302-3 +	6.00	86.13	2.4e+03	800
/cgn2_6/p/odata/2/1na/5B_CONB.seq:US-09-030-018-14 +	6.00	86.06	2.5e+03	807
/cgn2_6/p/odata/2/1na/5B_CONB.seq:US-08-276-192-1 +	6.00	86.06	2.5e+03	813
/cgn2_6/p/odata/2/1na/5C_CONB.seq:US-08-765-192-1 -	6.00	86.00	2.5e+03	817
/cgn2_6/p/odata/2/1na/5A_CONB.seq:US-08-776-088-12 -	6.00	86.00	2.5e+03	813
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-09145A-12 -	6.00	85.98	2.5e+03	816
/cgn2_6/p/odata/2/1na/6_CONB.seq:US-08-776-231-10 +	6.00	85.98	2.5e+03	816
/cgn2_6/p/odata/2/1na/5B_CONB.seq:US-08-330-161-3 +	6.00	85.95	2.5e+03	819
/cgn2_6/p/odata/2/1na/5B_CONB.seq:US-08-330-161-3 +	6.00	85.95	2.5e+03	819
/cgn2_6/p/odata/2/1na/5A_CONB.seq:US-07-941-523-1 +	6.00	85.92	2.5e+03	822
/cgn2_6/p/odata/2/1na/5B_CONB.seq:US-08-616-732A-8 +	6.00	85.92	2.5e+03	822
/cgn2_6/p/odata/2/1na/5B_CONB.seq:US-08-776-088-14 -	6.00	85.92	2.5e+03	822
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-09145A-14 -	6.00	85.92	2.5e+03	822
/cgn2_6/p/odata/2/1na/5D_CONB.seq:US-08-411-857-1 +	6.00	85.92	2.5e+03	822
/cgn2_6/p/odata/2/1na/5C_CONB.seq:US-08-790-137-2 +	6.00	85.89	2.5e+03	833
/cgn2_6/p/odata/2/1na/6_CONB.seq:US-08-733-230-3 -	6.00	85.73	2.6e+03	842
/cgn2_6/p/odata/2/1na/6_CONB.seq:US-08-630-457-6 +	6.00	85.51	2.6e+03	867
/cgn2_6/p/odata/2/1na/5C_CONB.seq:US-08-628-187-6 +	6.00	85.51	2.6e+03	867
/cgn2_6/p/odata/2/1na/5C_CONB.seq:US-08-744-026-2 +	6.00	85.47	2.7e+03	871
/cgn2_6/p/odata/2/1na/5D_CONB.seq:US-09-102-733-2 +	6.00	85.47	2.7e+03	871
/cgn2_6/p/odata/2/1na/5A_CONB.seq:US-08-457-745-6 +	6.00	85.45	2.7e+03	873
/cgn2_6/p/odata/2/1na/5B_CONB.seq:US-08-628-806B-1 +	6.00	85.42	2.7e+03	877
/cgn2_6/p/odata/2/1na/5B_CONB.seq:US-08-628-806B-1 +	6.00	85.38	2.7e+03	882
/cgn2_6/p/odata/2/1na/5D_CONB.seq:US-09-128-722-3 -	6.00	85.22	2.7e+03	900
/cgn2_6/p/odata/2/1na/5A_CONB.seq:US-08-460-481-6 -	6.00	85.16	2.8e+03	917
/cgn2_6/p/odata/2/1na/5B_CONB.seq:US-08-821-119-1 -	6.00	85.12	2.8e+03	917
/cgn2_6/p/odata/2/1na/5B_CONB.seq:US-08-764-100-22 +	6.00	85.12	2.8e+03	912
/cgn2_6/p/odata/2/1na/5B_CONB.seq:US-08-764-100-22 +	6.00	85.07	2.8e+03	918
/cgn2_6/p/odata/2/1na/5D_CONB.seq:US-08-833-659A-22 +	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/5A_CONB.seq:US-08-029-022-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/5B_CONB.seq:US-08-029-022-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/5B_CONB.seq:US-08-246-376-2 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/5C_CONB.seq:US-07-972-135-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/5C_CONB.seq:US-07-972-135-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/6_CONB.seq:US-08-256-004-2 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-2 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.97	2.8e+03	930
/cgn2_6/p/odata/2/1na/PCBUS_CONB.seq:PT-US99-05366-4 -	6.00	84.9		

/cgn2.6/prodata/2/1na/5B_COMB.seq:US-08-481-130-24 +	6.00	84.47	3.0e+03	992	1	/cgn2.6/prodata/2/1na/5D_COMB.seq:US-08-471-102-38 -	6.00	83.22	3.5e+03	1167	1
/cgn2.6/prodata/2/1na/5B_COMB.seq:US-08-656-984-24 +	6.00	84.47	3.0e+03	992	1	/cgn2.6/prodata/2/1na/5D_COMB.seq:US-08-471-102-38 -	6.00	83.21	3.6e+03	1168	1
/cgn2.6/prodata/2/1na/5B_COMB.seq:US-08-485-604-24 +	6.00	84.47	3.0e+03	992	1	/cgn2.6/prodata/2/1na/5D_COMB.seq:US-09-022-661-6 -	6.00	83.19	3.6e+03	1172	1
/cgn2.6/prodata/2/1na/5B_COMB.seq:US-08-487-595-24 +	6.00	84.47	3.0e+03	992	1	/cgn2.6/prodata/2/1na/5D_COMB.seq:US-09-022-661-8 -	6.00	83.19	3.6e+03	1172	1
/cgn2.6/prodata/2/1na/5B_COMB.seq:US-08-136-993-11 -	6.00	84.36	3.1e+03	993	1	/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-905-445-4 -	6.00	83.07	3.6e+03	1190	1
/cgn2.6/prodata/2/1na/5B_COMB.seq:US-08-285-641-11 -	6.00	84.36	3.1e+03	1006	1	/cgn2.6/prodata/2/1na/6_COMB.seq:US-08-955-625-4 -	6.00	83.07	3.6e+03	1190	1
/cgn2.6/prodata/2/1na/5D_COMB.seq:US-08-441-857-11 +	6.00	84.28	3.1e+03	1017	1	/cgn2.6/prodata/2/1na/6_COMB.seq:US-09-008-466-4 -	6.00	83.07	3.6e+03	1190	1
/cgn2.6/prodata/2/1na/5D_COMB.seq:US-08-441-857-11 +	6.00	84.28	3.1e+03	1017	1	/cgn2.6/prodata/2/1na/6_COMB.seq:US-09-008-466-4 -	6.00	83.07	3.6e+03	1190	1
/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-136-277-23 +	6.00	84.22	3.1e+03	1025	1	/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-680-957-7 -	6.00	82.97	3.7e+03	1205	1
/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-479-403-23 +	6.00	84.22	3.1e+03	1025	1	/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-403-852D-4 -	6.00	82.95	3.7e+03	1208	1
/cgn2.6/prodata/2/1na/6_COMB.seq:US-08-935-734-23 +	6.00	84.22	3.1e+03	1025	1	/cgn2.6/prodata/2/1na/5D_COMB.seq:US-08-358-171-9 -	6.00	83.00	3.7e+03	1201	1
/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-751-767-1 +	6.00	84.21	3.1e+03	1026	1	/cgn2.6/prodata/2/1na/6_COMB.seq:US-09-090-947-9 -	6.00	83.00	3.7e+03	1201	1
/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-807-050-2 -	6.00	84.16	3.1e+03	1033	1	/cgn2.6/prodata/2/1na/5B_COMB.seq:US-08-628-921-11 -	6.00	82.98	3.7e+03	1204	1
/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-891-254-8 -	6.00	84.14	3.2e+03	1035	1	/cgn2.6/prodata/2/1na/5D_COMB.seq:US-09-128-722-11 -	6.00	82.98	3.7e+03	1204	1
/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-819-539-8 -	6.00	84.14	3.2e+03	1035	1	/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-680-957-7 -	6.00	82.97	3.7e+03	1205	1
/cgn2.6/prodata/2/1na/5D_COMB.seq:US-09-030-270-8 -	6.00	84.14	3.2e+03	1035	1	/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-403-852D-4 -	6.00	82.95	3.7e+03	1208	1
/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-455-265-29 +	6.00	83.93	3.2e+03	1064	1	/cgn2.6/prodata/2/1na/5B_COMB.seq:US-08-553-703A-9 +	6.00	82.95	3.7e+03	1209	1
/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-449-043-29 +	6.00	83.93	3.2e+03	1064	1	/cgn2.6/prodata/2/1na/5B_COMB.seq:US-08-553-703A-9 +	6.00	82.95	3.7e+03	1209	1
/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-449-043-29 +	6.00	83.93	3.2e+03	1064	1	/cgn2.6/prodata/2/1na/5D_COMB.seq:US-09-006-021-5 +	6.00	82.95	3.7e+03	1209	1
/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-449-043-29 +	6.00	83.93	3.2e+03	1064	1	/cgn2.6/prodata/2/1na/5D_COMB.seq:US-09-006-021-5 +	6.00	82.95	3.7e+03	1209	1
/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-449-043-29 +	6.00	83.93	3.2e+03	1064	1	/cgn2.6/prodata/2/1na/5C_COMB.seq:US-09-006-021-5 +	6.00	82.95	3.7e+03	1209	1
/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-449-043-29 +	6.00	83.93	3.2e+03	1064	1	/cgn2.6/prodata/2/1na/5C_COMB.seq:US-09-006-021-5 +	6.00	82.95	3.7e+03	1209	1
/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-449-043-29 +	6.00	83.93	3.2e+03	1064	1	/cgn2.6/prodata/2/1na/5C_COMB.seq:US-09-006-021-5 +	6.00	82.95	3.7e+03	1209	1
/cgn2.6/prodata/2/1na/5C_COMB.seq:US-08-449-043-29 +	6.00	83.93	3.2e+03	1							

cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-08-046-431A-1-	6.00	81.84	4.2e+03	1395	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-08-123-161A-11-	6.00	81.84	4.2e+03	1396	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-08-483-278-11-	6.00	81.84	4.2e+03	1396	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-07-737-736B-5-	6.00	81.82	4.3e+03	1399	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-08-151-574-33-	6.00	81.79	4.3e+03	1404	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-08-146-424-19+	6.00	81.79	4.3e+03	1404	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-08-693-709-1+	6.00	81.79	4.3e+03	1404	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-08-487-111D-119-	6.00	81.79	4.3e+03	1404	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-419-448-33+	6.00	81.79	4.3e+03	1404	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-720-420A-119-	6.00	81.75	4.3e+03	1412	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-948-176-23+	6.00	81.68	4.3e+03	1426	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-358-171-3-	6.00	81.68	4.3e+03	1426	1
cgml2_6/p/oddata/2/1na/6_COMB.seg:US-09-030-947-3-	6.00	81.66	4.3e+03	1452	1
cgml2_6/p/oddata/2/1na/6_COMB.seg:US-09-027-166-8-	6.00	81.54	4.4e+03	1452	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-484-938B-23+	6.00	81.51	4.4e+03	1457	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-484-158B-23+	6.00	81.51	4.4e+03	1457	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-484-596A-23+	6.00	81.51	4.4e+03	1457	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-484-150A-23+	6.00	81.51	4.4e+03	1457	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-458-731-23+	6.00	81.51	4.4e+03	1457	1
cgml2_6/p/oddata/2/1na/6_COMB.seg:US-08-449-223A-23+	6.00	81.51	4.4e+03	1457	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-08-333-565-1+	6.00	81.43	4.5e+03	1472	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-08-191-337-2-	6.00	81.43	4.5e+03	1472	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-661-479-1+	6.00	81.43	4.5e+03	1472	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-644-271-31-	6.00	81.39	4.5e+03	1472	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-08-341-916-1+	6.00	81.38	4.5e+03	1482	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-08-805-166-1+	6.00	81.38	4.5e+03	1482	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-08-805-169-1+	6.00	81.38	4.5e+03	1482	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-957-365-1+	6.00	81.38	4.5e+03	1482	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-957-364-1+	6.00	81.38	4.5e+03	1482	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-941-647A-4+	6.00	81.33	4.5e+03	1491	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-310-912A-105+	6.00	81.33	4.5e+03	1491	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-310-912A-105+	6.00	81.33	4.5e+03	1491	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-820-170A-9+	6.00	81.32	4.5e+03	1493	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-632-470-49-	6.00	81.32	4.5e+03	1494	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-632-470-51-	6.00	81.30	4.5e+03	1498	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-324-977-3-	6.00	81.29	4.5e+03	1499	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-384-616-3-	6.00	81.29	4.5e+03	1499	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-07-914-281-3+	6.00	81.29	4.5e+03	1500	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-08-593-248-3+	6.00	81.29	4.5e+03	1500	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-08-525-058A-3+	6.00	81.29	4.5e+03	1500	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-696-731-3+	6.00	81.29	4.5e+03	1500	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-696-731-3+	6.00	81.29	4.5e+03	1500	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-08-716-301-9+	6.00	81.23	4.6e+03	1511	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-07-795-859B-25+	6.00	81.17	4.6e+03	1523	1
cgml2_6/p/oddata/2/1na/5A_COMB.seg:US-08-457-615-25+	6.00	81.17	4.6e+03	1523	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-948-176-24+	6.00	81.13	4.6e+03	1531	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-867-057-2+	6.00	81.06	4.7e+03	1544	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-128-369-2+	6.00	81.06	4.7e+03	1544	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-1-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-5-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-9-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-8-	6.00	81.06	4.7e+03	1545	1
cgml2_6/p/oddata/2/1na/5C_COMB.seg:US-08-628-039-7-	6.00	81.0			

[illegible]

[illegible]

/cgn2_6/ptodata/2/1na/5D_COMB.seq:US-08-839-032A-13 - 6.00 75.39 9.6e+03 3225
/cgn2_6/ptodata/2/1na/5D_COMB.seq:US-08-839-031A-13 - 6.00 75.39 9.6e+03 3225
/cgn2_6/ptodata/2/1na/5B_COMB.seq:PCT-US95-12724-13 - 6.00 75.33 9.7e+03 3251
/cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-346-455B-68 + 6.00 75.33 9.7e+03 3251
/cgn2_6/ptodata/2/1na/5B_COMB.seq:PCT-US95-06613-68 + 6.00 75.33 9.7e+03 3251
/cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-633-768A-3 + 6.00 75.29 9.7e+03 3288
/cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-208-008C-3 - 6.00 75.23 9.8e+03 3293
/cgn2_6/ptodata/2/1na/5D_COMB.seq:US-08-442-809A-75 - 6.00 75.21 9.8e+03 3300
/cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-194-290-6 - 6.00 75.21 9.8e+03 3300
/cgn2_6/ptodata/2/1na/5D_COMB.seq:US-08-928-592-29 - 6.00 75.21 9.8e+03 3300
/cgn2_6/ptodata/2/1na/5D_COMB.seq:US-08-614-377A-6 - 6.00 75.21 9.8e+03 3300
/cgn2_6/ptodata/2/1na/5C_COMB.seq:US-08-820-170A-33 - 6.00 75.16 9.9e+03 3324

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-325-553-1

seq_documentation_block:
; Sequence 1, Application US/08325553
; Patent No. 5538866
; GENERAL INFORMATION:
; APPLICANT: Israel, Ron S.
; APPLICANT: Heston, Warren D.W.
; TITLE OF INVENTION: THE PROSTATE-SPECIFIC MEMBRANE ANTIGEN
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325.553
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/973.337A
; FILING DATE: 05 NOV 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 1747/41426
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELETYPE: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2653 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Carcinoma
; IMMEDIATE SOURCE:
; CLONE: Prostate-Specific Membrane Antigen
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 262..2511
; US-08-325-553-1

alignment_scores: 9.00 Length: 9
Quality: 9

Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-080-127-2 x US-08-325-553-1 ..

Align seg 1/1 to: US-08-325-553-1 from: 1 to: 2653

291 AlaGluGluPheGlyLeuLeuGlySer 299

1528 GCAGAGAAATTGGCTCTTGCTTCT 1554

seq_name: /cgn2_6/ptodata/2/1na/5D_COMB.seq:US-08-394-152A-1

seq_documentation_block:
; Sequence 1, Application US/08394152A
; Patent No. 5935818
; GENERAL INFORMATION:
; APPLICANT: Israel, Ron S.
; APPLICANT: Heston, Warren D.W.
; TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM 330 466 DX2
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/394.152A
; FILING DATE: 24-FEB-95
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 41426-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2653 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Carcinoma
; IMMEDIATE SOURCE:
; CLONE: Prostate-Specific Membrane Antigen
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 262..2511
; US-08-394-152A-1

alignment_scores: 9.00 Length: 9
Quality: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-080-127-2 x US-08-394-152A-1 ..

Align seg 1/1 to: US-08-394-152A-1 from: 1 to: 2653

291 AlaGluGluPheGlyLeuGlySer 299
|||||
1528 GCGAGAGATTGTCTCTTGSTTCT 1554

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-221-817-7

seq_documentation_block:

; Sequence 7, Application US/08221817
; Patent No. 5532151
; GENERAL INFORMATION:
; APPLICANT: Chantry, David
; APPLICANT: Gray, Patrick W.
; APPLICANT: Hoekstra, Merle F.
; TITLE OF INVENTION: A No. 5532151el G Protein-Coupled Receptor
; TITLE OF INVENTION: Kinase GRK6
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221,817
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/123,932
; FILING DATE: 17 SEP 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5532151and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31981
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-221-817-7

alignment_scores:

Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-080-127-2 x US-08-221-817-7 ..

Align seg 1/1 to: US-08-221-817-7 from: 1 to: 32

117 AlaAspValAlaValValLys 123
|||||
12 GCGAGAGTGGCGCTGTCAG 32

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-221-817-8

seq_documentation_block:

; Sequence 8, Application US/08221817
; Patent No. 5532151
; GENERAL INFORMATION:
; APPLICANT: Chantry, David
; APPLICANT: Gray, Patrick W.
; APPLICANT: Hoekstra, Merle F.
; TITLE OF INVENTION: A No. 5532151el G Protein-Coupled Receptor
; TITLE OF INVENTION: Kinase GRK6
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221,817
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/123,932
; FILING DATE: 17 SEP 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5532151and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31981
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-221-817-8

alignment_scores:

Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-080-127-2 x US-08-221-817-8 ..

Align seg 1/1 to: US-08-221-817-8 from: 1 to: 32

117 AlaAspValAlaValValLys 123
|||||
12 GCGAGAGTGGCGCTGTCAG 32

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-454-439-7

seq_documentation_block:

; Sequence 7, Application US/08454439
; Patent No. 5591618
; GENERAL INFORMATION:
; APPLICANT: Chantry, David
; APPLICANT: Gray, Patrick W.
; APPLICANT: Hoekstra, Merle F.
; TITLE OF INVENTION: A No. 5591618el G Protein-Coupled Receptor

TITLE OF INVENTION: Kinase GRK6
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,439
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,817
FILING DATE: 31-MAR-1994
APPLICATION NUMBER: 08/123,932
FILING DATE: 17 SEP 1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: NO. 5591618and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31981
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-454-439-7

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-080-127-2 x US-08-454-439-7 ..

Align seg 1/1 to: US-08-454-439-7 from: 1 to: 32

117 Alaaspvalalavalvallys 123
|||||
12 GCGGACGTGGCGGTGTCACAG 32

seq_name: /cgn2_6/ptodata/2/1na/5a_COMB.seq:US-08-454-439-8
seq_documentation_block:
Sequence 8, Application US/08454439
Patent No. 5591618
GENERAL INFORMATION:
APPLICANT: Chantry, David
APPLICANT: Gray, Patrick W.
APPLICANT: Hoekstra, Merle F.
TITLE OF INVENTION: A No. 5591618el G Protein-Coupled Receptor
TITLE OF INVENTION: Kinase GRK6
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,439
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,817
FILING DATE: 31-MAR-1994
APPLICATION NUMBER: 08/123,932
FILING DATE: 17 SEP 1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: NO. 5591618and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31981
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-454-439-8

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-080-127-2 x US-08-454-439-8 ..

Align seg 1/1 to: US-08-454-439-8 from: 1 to: 32

117 Alaaspvalalavalvallys 123
|||||
12 GCGGACGTGGCGGTGTCACAG 32

seq_name: /cgn2_6/ptodata/2/1na/PCTUS_COMB.seq:PCT-US94-10487-7
seq_documentation_block:
Sequence 7, Application PC/TUS9410487
GENERAL INFORMATION:
APPLICANT: ICOS Corporation
TITLE OF INVENTION: A Novel G Protein-Coupled Receptor
TITLE OF INVENTION: Kinase GRK6
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10487
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/221,817
FILING DATE: 31 MAR 1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/123,932
FILING DATE: 17 SEP 1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Noland, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/31981
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3836
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
PCT-US94-10487-7

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-080-127-2 x PCT-US94-10487-7 ..

Align seq 1/1 to: PCT-US94-10487-7 from: 1 to: 32

117 AlaSPYAlaValVallys 123
|||||
12 GCGGACCTGGCGCTGTCAG 32

seq_name: /cgn2_6/ptodata/2/1na/PCTUS_COMB.seq: PCT-US94-10487-8

seq_documentation_block:
; Sequence 8, Application PC/TUS9410487
; GENERAL INFORMATION:
; APPLICANT: ICOS Corporation
; TITLE OF INVENTION: A Novel G Protein-Coupled Receptor
; TITLE OF INVENTION: Kinase GRK6
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10487
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/221,817
; FILING DATE: 31 MAR 1994
```

```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/123,932
FILING DATE: 17 SEP 1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Noland, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/31981
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3836
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
PCT-US94-10487-8

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-080-127-2 x PCT-US94-10487-8 ..

Align seq 1/1 to: PCT-US94-10487-8 from: 1 to: 32

117 AlaSPYAlaValVallys 123
|||||
12 GCGGACCTGGCGCTGTCAG 32

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq: US-08-411-389-19

seq_documentation_block:
; Sequence 19, Application US/08411389
; Patent No. 5605799
; GENERAL INFORMATION:
; APPLICANT: White, Raymond L.
; APPLICANT: Cawthon, Richard W.
; APPLICANT: Li, Ying
; TITLE OF INVENTION: SOMATIC MUTATIONS IN THE
; TITLE OF INVENTION: NEUROFIBROMATOSIS TYPE 1 GENE IN HUMAN TUMORS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue NW, Suite 1000
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/411,389
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/047,088
; FILING DATE: 16-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Innien, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 19780-107116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
```

```

: TELEFAX: 202-962-8300
: INFORMATION FOR SEQ ID NO: 19:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 62 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
:   MOLECULE TYPE: DNA (genomic)
:   HYPOTHETICAL: NO
:   ANTI-SENSE: NO
:   ORIGINAL SOURCE:
:   ORGANISM: Homo sapiens
: US-08-411-389-19

Alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

Alignment_block:
US-09-080-127-2 x US-08-411-389-19 ..

Align seg 1/1 to: US-08-411-389-19 from: 1 to: 62

482 AAlHsvalserGlyserGly 488
|||||
15 GCTCATGTTCTGCTGCTGCT 35

seq_name: /cgn2_6/ptodata/2/1na/backfile1.seq:5221620-16
seq_documentation_block:
: Patent NO. 5221620
: APPLICANT: PIRCHIO, ANTHONY F.; MADISEN, LINDA; WEBB, NANCY
: TITLE OF INVENTION: CLONING AND EXPRESSION OF TRANSFORMING
: GROWTH FACTOR BETA-2
: NUMBER OF SEQUENCES: 16
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US-07/446,020
:   FILING DATE: 05-DEC-1989
:   PRIOR APPLICATION DATA:
:     APPLICATION NUMBER: 285,140
:     FILING DATE: 16-DEC-1988
:     APPLICATION NUMBER: 234,065
:     FILING DATE: 18-AUG-1988
:     APPLICATION NUMBER: 148,267
:     FILING DATE: 25-JAN-1988
:     APPLICATION NUMBER: 106,752
:     FILING DATE: 06-OCT-1987
: SEQ ID NO: 16
: LENGTH: 116
: 5221620-16

Alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

Alignment_block:
US-09-080-127-2 x 5221620-16/rev ..

Align seg 1/1 to reverse of: 5221620-16 from: 1 to: 116

461 SerSerLeuHsArgArgAla 467
|||||
100 AGCTCCCTTCACCGCCGAGCT 80

seq_name: /cgn2_6/ptodata/2/1na/5D_COMB.seq:US-08-967-101-54
seq_documentation_block:
: Sequence 54, Application US/08967101
: Patent No. 5840540

: GENERAL INFORMATION:
: APPLICANT: ST. GEORGE-HYSLOP, PETER H
: APPLICANT: ROMMENS, JOHANNA M
: APPLICANT: FRASER, PAUL E
: TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
: TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
: NUMBER OF SEQUENCES: 183
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: TESTA, HURWITZ & THIBEAULT
: STREET: High Street Tower - 125 High Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: U.S.A.
: ZIP: 02110
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/967,101
: FILING DATE: 10-NOV-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/592,541
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Pitcher, Edmund R.
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 248-7000
: TELEFAX: (617) 248-7100
: INFORMATION FOR SEQ ID NO: 54:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 521 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-08-967-101-54

Alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

Alignment_block:
US-09-080-127-2 x US-08-967-101-54/rev ..

Align seg 1/1 to reverse of: US-08-967-101-54 from: 1 to: 521

294 PheGlyLeuLeuGlySerAsn 300
|||||
473 TTGGACTCTCTGGGCTCCAC 453

seq_name: /cgn2_6/ptodata/2/1na/5D_COMB.seq:US-08-592-541-54
seq_documentation_block:
: Sequence 54, Application US/08592541
: Patent No. 5986054
: GENERAL INFORMATION:
: APPLICANT: ST. GEORGE-HYSLOP, PETER H
: APPLICANT: ROMMENS, JOHANNA M
: APPLICANT: FRASER, PAUL E
: TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
: TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
: NUMBER OF SEQUENCES: 183
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: TESTA, HURWITZ & THIBEAULT
: STREET: High Street Tower - 125 High Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: U.S.A.
```

```

;
; ZIP: 02110
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7100
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-592-541-54

alignment_scores:
      quality: 7.00      length: 7
      ratio: 1.000      gaps: 0
      percent similarity: 100.000      percent identity: 100.000

alignment_block:
US-09-080-127-2 x US-08-592-541-54/rev ..

Align seq 1/1 to reverse of: US-08-592-541-54 from: 1 to: 521

294 PheGlyLeuLeuGlySerAsn 300
|||||
473 TTGGACTCTGGGCTCCAC 453

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-685-660A-4

seq_documentation_block:
; Sequence 4, Application US/08685660A
; Patent No. 5731412
; GENERAL INFORMATION:
; APPLICANT: SHIMOMURA, Takeshi
; APPLICANT: KAWAGUCHI, Toshiya
; APPLICANT: KITAMURA, Naomi
; TITLE OF INVENTION: NOVEL PROTEIN, DNA CODING FOR SAME
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 24-JUL-1996
; APPLICATION NUMBER: US/08/685,660A
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JPA Hei 7-187134
; FILING DATE: 24-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KIT, Gordon
; REGISTRATION NUMBER: 30,764
; REFERENCE/DOCKET NUMBER: Q-42295
```

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 759 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; STRAIN: MKN45
; FEATURE:
; NAME/KEY: coding sequence
; LOCATION: 1 to 759
; IDENTIFICATION METHOD: by experiment
; NAME/KEY: signal peptide
; LOCATION: 1 to 81
; IDENTIFICATION METHOD: by experiment
; NAME/KEY: mature peptide
; LOCATION: 82 to 759
; IDENTIFICATION METHOD: by experiment
; US-08-685-660A-4

alignment_scores:
      quality: 7.00      length: 7
      ratio: 1.000      gaps: 0
      percent similarity: 100.000      percent identity: 100.000

alignment_block:
US-09-080-127-2 x US-08-685-660A-4 ..

Align seq 1/1 to: US-08-685-660A-4 from: 1 to: 759

9 LeuLeuSerGlyValLeuAla 15
|||||
61 CTCCTCTCTGGGCTCGCGG 81

seq_name: /cgn2_6/ptodata/2/lna/5C_COMB.seq:US-08-974-196-4

seq_documentation_block:
; Sequence 4, Application US/08974196
; Patent No. 5854396
; GENERAL INFORMATION:
; APPLICANT: SHIMOMURA, Takeshi
; APPLICANT: KAWAGUCHI, Toshiya
; APPLICANT: KITAMURA, Naomi
; TITLE OF INVENTION: NOVEL PROTEIN, DNA CODING FOR SAME
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/974,196
; APPLICATION NUMBER:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/685,660
; FILING DATE: 24-JUL-1996
; APPLICATION NUMBER: JPA Hei 7-187134
; FILING DATE: 24-JUL-1995
```

```

ATTORNEY/AGENT INFORMATION:
NAME: KIT, Gordon
REGISTRATION NUMBER: 30,764
REFERENCE/DOCKET NUMBER: Q-42295
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 759 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna to mRNA
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
STRAIN: MKN45
FEATURE:
NAME/KEY: coding sequence
LOCATION: 1 to 759
IDENTIFICATION METHOD: by experiment
NAME/KEY: signal peptide
LOCATION: 1 to 81
IDENTIFICATION METHOD: by experiment
NAME/KEY: mature peptide
LOCATION: 82 to 759
IDENTIFICATION METHOD: by experiment
US-08-974-196-4

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-080-127-2 x US-08-974-196-4 ..
Align seg 1/1 to: US-08-974-196-4 from: 1 to: 759

9 LeuLeuSerglyValLeuAla 15
|||||
61 CTCCTCTCGGGCTCTGGCG 81

seq_name: /cgm2_6/ptodata/2/1na/PCTUS_COMB.seq:PCT-US91-01327-11

seq_documentation_block:
; Sequence 11, Application PC/TUS9101327
; GENERAL INFORMATION:
; APPLICANT: Draper et al.
; TITLE OF INVENTION: Oligonucleotide Therapies for
; TITLE OF INVENTION: Modulating the Effects of Herpesviruses
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Woodcock Washburn Kurtz
; ADDRESSEE: Mackiewicz & Norris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/01327
; FILING DATE: 19910225
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 485,297

```

```

FILING DATE: February 26, 1990
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: ISIS-0085
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1557
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: unknown
PCT-US91-01327-11

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-080-127-2 x PCT-US91-01327-11/rev ..
Align seg 1/1 to reverse of: PCT-US91-01327-11 from: 1 to: 1557

206 LeuAlaGluAlaGlySerVal 212
|||||
213 CTCGGGAGCGCGGATCGGTG 193

```


OM of: US-09-080-127-2 to: EST: * out_format : pfs
 Date: May 10, 2000 8:59 PM
 About: Results were produced by the Gencore software, version 4.5,
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=framed.p2n.model -DEV=xip
 -O=gn2_1/USPTO.spool/US09080127/runat.05052000_152835_6185/app-query.fasta.1
 -DB=EST -QMT=fastap -SUFFIX=first -GAPOP=4.500 -GAPEXT=0.050
 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=4.500
 -GAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000 -XGAPOP=6.000
 -DEEXT=7.000 -START=1 -MATRIX=oligo -TRANS=human40.cdi
 -LIST=LOCAL -DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15
 -MODE=LOCAL -OUTEXT=Dfs -NORM=ext -MILEN=0 -MAXLEN=1000000
 -USER=US09080127 -NCPU=6 -LOPU=3 -NO_XLPHY -WAIT -THREADS=1

Search information block:
 Query: US-09-080-127-2
 Query length: 496
 Database sequences: 4857316
 Database length: 202611650
 Search time (sec): 574.080000

WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPOP=XGAPOP=60.000
 WARN: XGAPEXT and YGAPEXT must be equal. Assuming XGAPEXT=YGAPEXT=60.000

Sequence	Strd Orig	ZScore	Escore	len	Document
gb_est2:AI213187	10.00	161.21	3.49	492	AI213187 y8d01a1.r1 Aspergillus
gb_est4:AM981869	10.00	158.83	4.73	682	AI3984169 EST297932 L. pennellii
gb_est34:AI188831	9.00	147.16	21.13	278	AI788531 uk46b04.r1 Soares mous
gb_est3:R35464	9.00	144.64	22.17	393	R35464 y8b0602.r1 Soares placet
gb_est14:AO988441	9.00	143.63	33.23	452	AO988441 HS_5088_B1_F04.SP6E RH
gb_est136:AI192895	9.00	143.52	33.71	459	AI192895 mg24h05.y1 Barstead MR
gb_est46:AM287578	9.00	143.42	34.12	465	AM287578 LG1_242_H05.D1_A02C.L4
gb_gss9:AO079756	9.00	143.39	34.25	467	AO079756 CIR-HSP-235810.TR CIT
gb_gss1:CN0500UM	9.00	142.67	37.59	516	AI091612 Arabidopsis thaliana R
gb_gss14:AO431348	9.00	142.47	38.54	530	AO431348 HS_5053_A1_E10.SP6E RH
gb_gss5:AO777636	9.00	142.10	40.43	558	AO777636 HS_2252_B2_F04.MR CIT
gb_gss12:AO274419	9.00	140.86	47.41	662	AO274419 mg24h05.y1 Barstead MR
gb_est10:AI152553	8.00	134.54	102.72	144	AI152553 m85f03.r1 Stratagene
gb_est19:AA790684	8.00	133.54	121.12	148	AA790684 v418c08.r1 Soares_mamm
gb_est43:AV370653	8.00	133.40	123.40	151	AV370653 AV370653 RIKEN full-16
gb_gss1:FR0010068	8.00	132.41	140.07	173	AI001346 F.rubripes GSS sequenc
gb_est9:AA078132	8.00	131.29	151.83	202	AA078132 TH14E11 Chromosome 7 H
gb_est23:AI115705	8.00	129.85	194.43	246	AI115705 uc34d03.r1 Soares_mamm
gb_est14:AA82648	8.00	129.37	206.92	263	AA82648 EST96167 Testis I Homc
gb_est24:AI163892	8.00	129.05	215.70	275	AI163892 AO51P05U Hyb1d aspen
gb_est7:W61394	8.00	128.32	226.81	304	W61394 md92b01.r1 Soares mous
gb_gss16:AO558944	8.00	128.29	227.54	305	AO558944 Sheared DNA-27H20.TR S
gb_est30:AB020055	8.00	128.11	243.34	313	AB020055 AB020055 Riken full-16
gb_est33:AV121865	8.00	128.08	244.06	314	AV121865 AV21865 Riken full-16
gb_est29:AI108665	8.00	127.92	249.13	321	AI108665 tw18f04.r1 NCI.CGAP.BH
gb_est11:AA254486	8.00	127.72	255.63	330	AA254486 v417b04.r1 Soares mous
gb_est28:AI165993	8.00	127.72	255.63	330	AI165993 v418c08.r1 Soares_mamm
gb_est12:R00815	8.00	127.68	257.07	332	FO0815 HSB66A072 STRATAGENE Hunt
gb_est7:W61392	8.00	127.48	263.56	341	W61392 md92b02.r1 Soares mous
gb_est23:AI122169	8.00	127.42	265.72	344	AI122169 udi13h08.r1 Soares thym
gb_est4:R35188	8.00	127.42	266.43	345	H35188 EST11862 Rat PC-12 cell
gb_est30:AI162410	8.00	127.29	270.03	350	AI162410 vs97d02.r1 Barstead mc
gb_gss9:AO026731	8.00	127.25	271.47	352	AO026731 CIR-HSP-232463.TR CIT
gb_est7:W58997	8.00	127.21	272.90	354	W58997 md74e01.r1 Soares mous
gb_gss13:AO345053	8.00	126.91	283.66	369	AO345053 RPI11-121J1.TV RPI1
gb_gss13:AO345066	8.00	126.77	288.67	376	AO345066 RPI11-121J1.TV RPI1
gb_est8:W81733	8.00	126.75	289.38	377	W81733 me94e06.r1 Soares mous
gb_est31:AI1676300	8.00	126.73	290.10	378	AI1676300 etmest0045 EHI Elmerl
gb_est7:W55714	8.00	126.70	291.53	380	W55714 md14e03.r1 Soares mous

gb_est9:AA100584	8.00	126.68	292.24	381	AA100584 zn51n06.r1 Striatage
gb_est8:AA002277	8.00	126.62	294.35	384	AA002277 md42b03.r1 Soares m
gb_est17:C72605	8.00	126.53	297.99	384	C72605 C72605 Rice panicle a
gb_gss12:AO278408	8.00	126.45	299.38	391	AO278408 CIRB1-E1-251441.TR
gb_est14:AA431164	8.00	126.45	300.81	393	AA431164 zw71609.r1 Soares.L
gb_est7:W58774	8.00	126.40	302.95	396	W58774 md19f09.r1 Soares mous
gb_est14:AA396829	8.00	126.38	303.66	397	AA396829 mw40a05.r1 Life Tec
gb_est41:AA158862	8.00	126.38	303.66	397	AA158862 za48e11.r1 Xenopus
gb_gss9:AO024256	8.00	126.34	305.08	399	AO024256 HS_2184_A1_E02.MR C
gb_est19:AA815574	8.00	126.34	308.64	404	AA815574 vs97d02.r1 Barstead
gb_est4:H26724	8.00	126.09	314.33	412	H26724 Y11412.S1 Soares bre
gb_est30:AI165473	8.00	126.04	315.04	413	AI165473 mr77a05.y1 Striatage
gb_est14:AA116321	8.00	126.04	317.17	416	AA116321 v601d12.S1 Knowles
gb_est9:AA073435	8.00	126.02	317.88	417	AA073435 mm93f11.r1 Striatage
gb_est7:W58570	8.00	126.02	320.72	421	W58570 zd21e10.S1 Soares.fet
gb_est27:AI130788	8.00	125.90	322.85	424	AI130788 mm67b05.y1 Soares m
gb_est19:AA765939	8.00	125.85	324.98	427	AA765939 ca14e11.S1 NCI.CGAP
gb_est45:AA291396	8.00	125.80	327.11	430	AA291396 UR-B12-agg-b-01.0
gb_gss8:AO276382	8.00	125.76	328.52	432	AO276382 RPI1-23-133014.TV R
gb_est20:AA836389	8.00	125.75	329.23	433	AA836389 0337c05.S1 NCI.CGAP
gb_est19:AA797101	8.00	125.63	334.19	440	W97024 m88a12.r1 Soares mou
gb_est12:W97024	8.00	125.63	334.19	441	W97024 m88a12.r1 Soares mou
gb_est29:AI1603095	8.00	125.53	338.43	446	AI1603095 UR-R-AG0-xc-g-08-0
gb_est23:AI098363	8.00	125.50	339.84	448	AI098363 v684e11.r1 Barstead
gb_est20:AA870803	8.00	125.44	342.67	452	AA870803 v626f09.r1 Barstead
gb_est12:W47827	8.00	125.32	347.61	459	D47827 R1C51353A Rice green
gb_est17:AA600392	8.00	125.29	349.02	461	AA600392 v55b11.r1 Barstead
gb_est4:R90684	8.00	125.28	349.72	462	R90684 17039 Lambda-PRL2 Ara
gb_est47:AO028331	8.00	125.25	351.13	464	AO028331 AO028331 POS10M1.P1
gb_est14:AA879533	8.00	125.21	352.54	466	AA879533 v687e09.r1 Ko mous
gb_est30:AA187763	8.00	125.21	352.54	466	AA187763 uk44d01.r1 Soares m
gb_est35:AI148799	8.00	125.21	352.54	466	AI148799 UR-M-AJ1-ah-e-10.0
gb_est13:AA028331	8.00	125.18	353.95	468	AA028331 v929f08.r1 Soares_m
gb_est10:AI157488	8.00	125.15	355.36	470	AI157488 yp60f11.S1 Soares.fet
gb_est4:H40466	8.00	125.15	355.36	470	H40466 yp60f11.S1 Soares.fet
gb_gss3:B38775	8.00	125.08	358.18	474	B38775 HS-1048-A2-D10-MR.ab1
gb_gss13:AO357990	8.00	125.08	358.18	474	AO357990 CIRB1-E1-2529121.TR
gb_est8:AA023048	8.00	125.08	358.18	475	AA023048 mh67b05.r1 Soares m
gb_gss12:AO282732	8.00	125.08	358.18	475	AO282732 RPI11-77G12.TV RPI
gb_gss10:AA440592	8.00	125.03	360.29	477	AA440592 vs94e07.r1 Striatage
gb_gss13:AO404333	8.00	125.03	360.29	478	AO404333 HS_5069_B1_F01.SP6E
gb_est31:AI167605	8.00	125.00	362.40	480	AI167605 etmest0158 Eth1 Elm
gb_est27:AI146696	8.00	124.91	366.62	486	AI146696 t198a05.y1 NCI.CGAP
gb_gss7:AO913704	8.00	124.85	369.43	490	AO913704 m8eb004B17 CIGI R
gb_est23:AI117488	8.00	124.82	371.53	493	AI117488 ub88e11.r1 Soares.L
gb_est12:AI159381	8.00	124.75	374.34	497	AI159381 vs97d02.r1 Barstead
gb_est29:AI159381	8.00	124.75	374.34	497	AI159381 vs97d02.r1 Barstead
gb_est46:AA272994	8.00	124.69	377.15	501	AA272994 MKO-RH0208-101299-2
gb_est35:AI151537	8.00	124.65	379.25	504	AI151537 UR-M-BH0-ak1-C-07-0
gb_gss1:FR0022616	8.00	124.62	380.65	506	FR0022616 m85f01.r1 Soares m
gb_est11:AA078132	8.00	124.60	381.35	507	AA078132 v18f11.r1 Soares K
gb_est15:AA521590	8.00	124.56	383.45	510	AA521590 v107f01.r1 Barstead
gb_est11:AA110375	8.00	124.56	384.15	511	AA110375 mh72c08.r1 Barstead
gb_est3:AI153542	8.00	124.55	384.15	511	AI153542 UR-M-BH0-aj3-a-04.0
gb_est28:AI108268	8.00	124.52	385.56	513	AI108268 mh67b05.y1 Soares m
gb_est20:AA663998	8.00	124.48	389.05	518	AA663998 v607e11.r1 Barstead
gb_gss15:AO421791	8.00	124.38	392.55	523	AO421791 RPI1-1-427M21.TJ R
gb_gss14:AO450945	8.00	124.31	396.05	528	AO450945 HS_5152_A2.S12.T7A
gb_est32:AI1757603	8.00	124.28	397.44	529	AI1757603 EESTeal8e03.y1 Elm
gb_est43:AI1757603	8.00	124.28	397.44	530	AI1757603 P11C-PK002.o12 chl
gb_gss12:AO316845	8.00	124.27	398.14	531	AO316845 RPI11-3H18.TJ RPI1
gb_est11:AA15011	8.00	124.24	399.54	533	AA15011 mh83f07.r1 Striatage
gb_est14:AA386853	8.00	124.24	399.54	533	AA386853 v621g03.r1 Ko mous
gb_gss16:AO399673	8.00	124.23	400.24	534	AO399673 HS_5368_B2_G08.T7A
gb_est3:AI1757566	8.00	124.23	400.24	534	AI1757566 EESTeal8a04.y1 Elm
gb_est12:AA85744	8.00	124.14	404.42	540	AA85744 v82c03.r1 Soares m
gb_est17:AA655398	8.00	124.02	410.70	549	AA655398 v10c11.r1 Striatage
gb_gss6:AO844419	8.00	124.02	410.70	549	AO844419 an37d06 JM101 file

gb_gss6:AO894408	+	8.00	124.01	411.40	550	AO894408	HS_3224.B1	B10.T7C	CIT	gb_est15:AA42961	-	7.00	115.30	1.3e+03	149	AA42961	v11a03.r1	Stratage	
gb_est11:AA222947	+	8.00	123.93	415.57	556	AA222947	zr13612.s1	Stratage		gb_est17:AI940084	+	7.00	115.25	1.3e+03	150	AI940084	RC2-CIT003-1	20799-0	
gb_est16:AW377983	+	8.00	123.92	416.27	557	AW377983	MR0-TI0208	-22199-204		gb_est17:AI940084	+	7.00	115.25	1.3e+03	150	AI940084	RC2-CIT003-1	20799-0	
gb_est13:AI155919	+	8.00	123.88	418.36	560	AI155919	ue01d09.r1	Soares.NMx		gb_est14:MM16482	+	7.00	115.21	1.3e+03	150	MM16482	RC2-CIT003-1	240899-0	
gb_est10:AI648008	+	8.00	123.76	424.62	569	AI648008	uk39p03.r1	Sugano.mous		gb_gss8:AV071296	+	7.00	115.21	1.3e+03	151	AV071296	RPC1-23	-331A14.TV	
gb_est17:AI430099	+	8.00	123.71	427.40	573	AI430099	md74e01.r1	Soares.mous		gb_est14:MM16482	+	7.00	115.16	1.3e+03	152	MM16482	RC2-CIT003-1	240899-0	
gb_est13:AO23084	+	8.00	123.66	430.18	577	AO23084	AO23084	Mouse.unferf		gb_est1:RM6648	+	7.00	115.06	1.3e+03	154	RM6648	y35510.s1	Soares.N	
gb_est18:AM055367	+	8.00	123.64	431.57	579	AM055367	614004H12.r1	614 - roc		gb_est10:AV222047	+	7.00	114.97	1.3e+03	156	AV222047	AV232047	RIKEN.fu1	
gb_est16:AA537555	+	8.00	123.51	438.50	589	AA537555	vx48e08.r1	Soares.mamm		gb_est18:AA676300	+	7.00	114.61	1.4e+03	164	AA676300	ad37e06.s1	Stratage	
gb_est14:AA445418	+	8.00	123.43	439.89	591	AA445418	v65e04.r1	Bairstead.MH		gb_est11:AV029200	+	7.00	114.61	1.4e+03	164	AV029200	AV029200	Mus.muscul	
gb_est10:AM737493	+	8.00	123.43	443.36	596	AM737493	MR0-TI0208	-101299-207		gb_est13:AV05054	+	7.00	114.56	1.4e+03	165	AV05054	Y619F04.s1	Soares.inf	
gb_est10:AM737493	+	8.00	123.39	445.44	599	AM737493	vg22b02.r1	Bairstead.st		gb_est14:AV132721	+	7.00	114.47	1.4e+03	167	AV132721	AV152721	Mus.muscul	
gb_gss13:AO395926	+	8.00	123.38	446.13	600	AO395926	CITB1-EI-2542p13	TF.CD		gb_est14:AV132025	+	7.00	114.43	1.4e+03	168	AV132025	AV152025	Mus.muscul	
gb_est13:AM809948	+	8.00	123.88	475.84	643	AM809948	t158c10.r1	Soares.NSL		gb_est1:TM4513	+	7.00	114.35	1.4e+03	170	TM4513	224508.r1	Stratage	
gb_est15:AM257514	+	8.00	122.72	485.48	657	AM257514	EST305651	KV2	Medicac	gb_est10:AA904332	+	7.00	114.30	1.4e+03	171	AA904332	oe/6911.s1	NCI.CGAP	
gb_est14:AM176598	+	8.00	122.70	486.86	659	AM176598	te02d10.r1	Zebrafish.W		gb_est15:AI870098	+	7.00	114.30	1.4e+03	171	AI870098	wm65c08.x1	NCI.CGAP	
gb_gss13:AO547186	+	8.00	122.69	487.55	660	AO547186	RPCI-11	-427M19.TV	RPCI	gb_gss8:BA8431	+	7.00	114.18	1.5e+03	174	BA8431	RPCI1-1M18	TV	
gb_gss13:AO547186	+	8.00	122.65	489.92	662	AO547186	CITB1-EI-2529122	TF.CD		gb_est13:AV028552	+	7.00	113.97	1.5e+03	179	AV028552	AV028552	RIKEN.fu1	
gb_est10:AM176640	+	8.00	122.65	489.92	663	AM176640	mt14f07.r1	Soares.mous		gb_est13:AA362627	+	7.00	113.77	1.5e+03	184	AA362627	ESR72638	Ovary	
gb_gss16:AO641774	+	8.00	122.52	497.86	675	AO641774	RPCI03-DPMT	-25N13.TV		gb_est15:AA526830	+	7.00	113.77	1.5e+03	184	AA526830	n152c07.s1	NCI.CGAP	
gb_est13:AM176840	+	8.00	122.21	518.44	705	AM176840	t133c09.r1	Sugano.Kava		gb_est10:AI658775	+	7.00	113.73	1.5e+03	185	AI658775	lu08e07.x1	NCI.CGAP	
gb_gss1:AG002887	+	8.00	121.91	538.28	734	AG002887	Homo	sapiens	genomic.D	gb_est13:AI805049	+	7.00	113.73	1.5e+03	185	AI805049	ts33c11.x1	NCI.CGAP	
gb_est10:AA05184	+	8.00	120.95	609.00	838	AA051784	m56e05.r1	Soares.mous		gb_est16:AI918467	+	7.00	113.73	1.5e+03	185	AI918467	ts31d05.x1	NCI.CGAP	
gb_gss8:AO900608	+	8.00	120.94	610.35	840	AO900608	HS_3225	B1.A04	TF.CIT	gb_est17:AI973049	+	7.00	113.65	1.6e+03	187	AI973049	w46h10.x1	NCI.CGAP	
gb_gss1:CN5004CD	+	8.00	120.16	673.74	934	AI051260	Drosophila	melanogast		gb_est12:AV325102	+	7.00	113.65	1.6e+03	187	AV325102	AV325102	RIKEN.fu1	
gb_gss1:CN5004YWS	+	8.00	119.96	691.86	961	AL056934	Drosophila	melanogast		gb_est16:NM5925	+	7.00	113.61	1.6e+03	188	NM5925	26-AD4	multi-tissue	
gb_gss1:CN5004YWS	+	8.00	119.76	709.28	987	AL057011	Drosophila	melanogast		gb_est19:AV026237	+	7.00	113.61	1.6e+03	188	AV026237	AV026237	RIKEN.fu1	
gb_gss1:CN5004YWS	+	8.00	119.67	717.98	1000	AL057011	Drosophila	melanogast		gb_est13:AV029547	+	7.00	113.58	1.6e+03	189	AV029547	AV029547	Mus.muscul	
gb_gss1:CN5004YWS	+	8.00	119.42	741.36	1035	AL056802	Drosophila	melanogast		gb_est14:AI201556	+	7.00	113.54	1.6e+03	190	AI201556	g974f06.x1	NCI.CGAP	
gb_gss1:CN5004YWS	+	8.00	118.97	785.30	1101	AL056880	Drosophila	melanogast		gb_est16:AI918467	+	7.00	113.54	1.6e+03	190	AI918467	ts31d05.x1	NCI.CGAP	
gb_gss1:CN5004YWS	+	8.00	118.71	811.85	1141	AO743323	HS_5388	B1.C03	SP6	RD	gb_est15:AI868905	+	7.00	113.54	1.6e+03	190	AI868905	w46h10.x1	NCI.CGAP
gb_gss1:AO743323	+	8.00	117.83	909.52	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est16:NM5925	+	7.00	113.31	1.6e+03	194	NM5925	26-AD4	multi-tissue
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388	B1.C03	SP6	RD	gb_est13:AI154503	+	7.00	113.31	1.6e+03	194	AI154503	u630h10.r1	Soares.m
gb_gss1:AO743323	+	8.00	117.61	926.03	1289	AO743323	HS_5388												

gb_est18:AA026490	7.00	112.02	1.9e+03	234	AA026490 z692903.r1 Soares_fetd	gb_est11:AA235532	7.00	111.07	2.2e+03	267	AA235737 z53767.r1 Soares
gb_est19:AA017931	7.00	112.02	1.9e+03	233	AA017931 CT17431 Human placenta cd	gb_gss6:AA0881786	7.00	111.07	2.2e+03	268	AA0881796 h5-5667.r1 P02_SpHb
gb_est12:AA220300	7.00	112.02	1.9e+03	233	AA220300 EST22737 Adipose_tlsu	gb_est4:AD61182	7.00	111.04	2.2e+03	267	AD61182 HUM210606b Clontech_h
gb_est16:AA570936	7.00	112.02	1.9e+03	234	AA570936 v185ell.r1 Stratagene	gb_est11:1928385	7.00	111.01	2.2e+03	269	1928385 EST153214 Human Embryo
gb_est127:AA1545921	7.00	112.02	1.9e+03	234	AA1545921 UI-R-CAP-gl-e-11-0-UI-	gb_est12:AA225809	7.00	111.01	2.2e+03	269	AA225809 EST228870 Cerebellum
gb_est33:AA18555281	7.00	112.02	1.9e+03	234	AA18555281 603013H03.x1 603 - str	gb_est16:AA548813	7.00	111.01	2.2e+03	269	AA548813 RK02c06.s1 NCI CGAR
gb_est43:AAV440518	7.00	112.02	1.9e+03	234	AAV440518 AV340518 RIKEN full-1e	gb_est16:AA192019	7.00	111.01	2.2e+03	269	AA192019 614022069.x1 614 -
gb_est41:AAV276349	7.00	111.96	1.9e+03	236	AAV276349 AV276349 RIKEN full-1e	gb_est16:AA29919	7.00	110.99	2.2e+03	270	AA29919 zw07410.r1 Soares_N
gb_est45:AAW915157	7.00	111.96	1.9e+03	236	AAW915157 UI-H-B12-egk-e-08-0-UI-	gb_est12:AA978123	7.00	110.99	2.2e+03	270	AA978123 0p32610.r1 Soares_N
gb_est34:AAV171030	7.00	111.93	1.9e+03	237	AAV171030 AV171030 Mus musculus	gb_est24:AA193458	7.00	110.99	2.2e+03	270	AA193458 EST231220 Normalized
gb_est27:AA150596	7.00	111.90	1.9e+03	238	AA150596 m29C04.x1 Stratagene	gb_est26:AA373117	7.00	110.99	2.2e+03	270	AA373117 A1237117 Emericella
gb_est41:AAV80859	7.00	111.90	1.9e+03	238	AAV808576 AV280576 RIKEN full-1e	gb_est34:AAV154887	7.00	110.99	2.2e+03	270	AAV154887 AV154987 Mus muscul
gb_gss3:AA831970	7.00	111.90	1.9e+03	238	B83176 RCTC11-25C13.TV RCTC-11	gb_est40:AAV419971	7.00	110.99	2.2e+03	270	AAV419971 KX318707.x1 NCI CGAR
gb_est16:AAV90786	7.00	111.87	2.0e+03	239	AAV90786 vma1C09.r1 Knowles Sol	gb_est19:AAV012129	7.00	110.96	2.2e+03	271	AA812129 0548712.s1 NCI CGAR
gb_est17:AA647662	7.00	111.84	2.0e+03	240	AA647662 vp03c08.r1 Soares_mamm	gb_est33:AAV012717	7.00	110.96	2.2e+03	271	AAV012717 AV102717 Mus muscul
gb_est43:AAV537818	7.00	111.84	2.0e+03	240	AAV537818 AV537818 RIKEN full-1e	gb_est34:AAV169963	7.00	110.96	2.2e+03	271	AAV16963 AV169963 Mus muscul
gb_est44:AAV376718	7.00	111.81	2.0e+03	241	AAV376718 AV376718 RIKEN full-1e	gb_est36:AAV1925705	7.00	110.96	2.2e+03	271	AAV1925705 w033309.x1 NCI CGAR
gb_est42:AAV633093	7.00	111.78	2.0e+03	242	AA633093 nc10a06.s1 NCI CGAR-TH	gb_est42:AAV27868	7.00	110.96	2.2e+03	271	AAV27868 AV327868 RIKEN full11
gb_est44:AAV287925	7.00	111.78	2.0e+03	242	AAV287925 AV287925 RIKEN full-1e	gb_est37:AAV98828	7.00	110.93	2.2e+03	272	AAV98828 mg05111.r1 Soares mo
gb_est44:AAV51566	7.00	111.78	2.0e+03	242	AAV51566 AV51566 RIKEN full-1e	gb_est34:AAV154898	7.00	110.93	2.2e+03	272	AAV154898 AV154898 Mus muscul
gb_est49:AAV515044	7.00	111.75	2.0e+03	243	AAV515045 ISSG0028 Rice Immature	gb_est46:AAW47463	7.00	110.91	2.2e+03	273	AAW47463 31214 MNC P2IG Susa7
gb_est43:AAV51186	7.00	111.75	2.0e+03	243	AAV51186 AV51186 RIKEN full-1e	gb_est26:AAV165097	7.00	110.88	2.2e+03	274	AAV165097 q956802.x1 NCI CGAR
gb_est43:AAV51186	7.00	111.75	2.0e+03	243	AAV51186 AV51186 RIKEN full-1e	gb_est27:AAV440763	7.00	110.88	2.2e+03	274	AAV440763 s653309.r1 Gm-c1004
gb_est22:AA1007915	7.00	111.72	2.0e+03	244	AA1007915 EST202366 Normalized	gb_est30:AAV159157	7.00	110.88	2.2e+03	274	AAV159157 tt85N10.x1 NCI CGAR
gb_est23:AA1553491	7.00	111.72	2.0e+03	244	AA155391 uc50b05.r1 Soares_thym	gb_est41:AAV70709	7.00	110.88	2.2e+03	274	AAV70709 AV270709 RIKEN full-1e
gb_est34:AA1804116	7.00	111.72	2.0e+03	244	AA1804116 tc68d11.x1 Soares_NHM	gb_gss10:AAQ098627	7.00	110.88	2.2e+03	274	AAQ098627 h8-3044.r1 B04_MRC
gb_est34:AAV124000	7.00	111.72	2.0e+03	244	AAV124000 AV124000 Mus musculus	gb_est34:AA1802492	7.00	110.85	2.2e+03	275	AA1802492 tp08604.x1 NCI CGAR
gb_est11:AA272757	7.00	111.72	2.0e+03	245	AA13033 HSC20G062 normalized inf	gb_est39:AAV129066	7.00	110.85	2.2e+03	275	AAV129066 UI-M-BH2.3-Cob-d-06050
gb_est44:AAV050319	7.00	111.66	2.0e+03	246	AAV050319 UI-H-B11-efit-g-05-0-UI-	gb_est42:AAV227547	7.00	110.85	2.2e+03	275	AAV227547 AV322794 RIKEN full-1e
gb_est24:AAV121254	7.00	111.66	2.0e+03	247	AA1212594 x8f05a1.r1 Aspergillus	gb_est11:AD4914	7.00	110.83	2.2e+03	276	AD4914 R1C827774 Rice root C
gb_est43:AAV559526	7.00	111.63	2.0e+03	247	AAV559526 AV559526 RIKEN full-1e	gb_est18:AAV109584	7.00	110.83	2.2e+03	276	AAV09258 A121208.s1 Soares_L
gb_est11:AA272757	7.00	111.60	2.0e+03	248	AA272757 vb09h12.r1 Soares_mous	gb_est46:AAW61329	7.00	110.83	2.2e+03	276	AAW61329 RC3-CT0254.r1 Soares-O
gb_est34:AAV173580	7.00	111.60	2.0e+03	248	AAV173580 AV173580 Mus musculus	gb_est31:AAV033009	7.00	110.80	2.2e+03	277	AAV033009 AV033009 Mus muscul
gb_est41:AAV257537	7.00	111.60	2.0e+03	248	AAV257537 AV257537 RIKEN full-1e	gb_est14:AA455084	7.00	110.77	2.2e+03	278	AA455084 aa04605.s1 Soares_N
gb_est41:AAV278524	7.00	111.60	2.0e+03	248	AAV278524 AV278524 RIKEN full-1e	gb_est34:AAV13408	7.00	110.77	2.2e+03	278	AAV13408 AV13408 Mus muscul
gb_est31:AAV041193	7.00	111.54	2.0e+03	250	R73683 y197607.r1 Soares_breast	gb_est12:AA306693	7.00	110.75	2.3e+03	279	AA306693 EST17767 Acta ended
gb_est34:AAV041193	7.00	111.54	2.0e+03	250	AAV041193 AV041193 Mus musculus	gb_est13:AA333156	7.00	110.75	2.3e+03	279	AA333156 EST371222 Embryo_8
gb_est44:AAV1570024	7.00	111.54	2.0e+03	250	AAV150087 AV150087 Mus musculus	gb_est15:AA472814	7.00	110.72	2.3e+03	280	AA472814 v058611.r1 Knowles
gb_est12:1966508	7.00	111.46	2.1e+03	253	R66056 vt39ef07.r1 Soares_infnt	gb_est16:AA687851	7.00	110.72	2.3e+03	280	AA687851 nv065012.s1 NCI CGAR
gb_est40:AAV454543	7.00	111.46	2.1e+03	253	AAV45453 AV245453 RIKEN full-1e	gb_est34:AAV146596	7.00	110.72	2.3e+03	280	AAV146596 AV207693 Mus muscul
gb_gss3:AAV093928	7.00	111.46	2.1e+03	253	AAV093928 AV093928 Mus musculus	gb_est10:AA618958	7.00	110.72	2.3e+03	280	AAV093928 AV207693 RIKEN full11
gb_gss7:AA0911485	7.00	111.26	2.1e+03	260	AA0844326 an36q07 JM101 flittered	gb_est11:AA234434	7.00	110.70	2.3e+03	281	AA234434 z772105.r1 Soares_N
gb_est17:AA0843826	7.00	111.26	2.1e+03	261	AA658171 nuf6001.s1 NCI CGAR_Pt	gb_est27:AA119631	7.00	110.70	2.3e+03	281	AA119631 t952101.x1 Soares_N
gb_est121:AAV000319	7.00	111.23	2.1e+03	261	AAV000319 AT000319 Apple carpel	gb_est34:AAV126931	7.00	110.70	2.3e+03	281	AAV126931 AV126931 Mus muscul
gb_est42:AAV000319	7.00	111.23	2.1e+03	261	AAV15058 AV15058 Mus musculus	gb_est37:AA1934342	7.00	110.70	2.3e+03	281	AA1934342 614029H12.x1 614 -
gb_est43:AAV15058	7.00	111.23	2.1e+03	261	AAV15058 AV15058 Mus musculus	gb_est22:AA1010291	7.00	110.67	2.3e+03	282	AA1010291 EST2204742 Normalized
gb_est43:AAV15058	7.00	111.23	2.1e+03	261	AAV15058 AV15058 Mus musculus	gb_est39:AAV115493	7.00	110.67	2.3e+03	282	AAV15493 vt67604.y1 Sommer P
gb_gss3:AA1710	7.00	111.23	2.1e+03	261	B31710 HS-1012-B1-A06-MR.ab1 CT	gb_est11:AA330554	7.00	110.64	2.3e+03	283	AA330554 v650603.r1 Stratagene
gb_est43:AAV172780	7.00	111.20	2.1e+03	262	AAV172780 UI-R-Y0-LW-a-03-0-UI-S	gb_est13:AA444941	7.00	110.64	2.3e+03	283	AA444941 EST508872 Gall1 bladd
gb_est40:AAV013447	7.00	111.12	2.1e+03	265	AAV013447 AV013447 Mus musculus	gb_est33:AAV214452	7.00	110.64	2.3e+03	283	AAV214452 AV214452 RIKEN full11
gb_est44:AAV1504472	7.00	111.12	2.1e+03	265	AAV1504472 xg52d02.x1 NCI CGAR_UT	gb_est16:AA616455	7.00	110.62	2.3e+03	284	AA616455 zF-E500 zebrafish ec
gb_est44:AAV1645195	7.00	111.12	2.1e+03	266	AAV1645195 84568 MARC_LBOV Bos ta	gb_est10:AA618958	7.00	110.57	2.3e+03	286	AAV16963 AV207693 RIKEN full11
gb_est330:AA1528652	7.00	111.09	2.2e+03	266	AA1528652 t777909.x1 NCI CGAR_K3	gb_est29:AA105855	7.00	110.44	2.3e+03	291	AA105855 vt28112.x1 Batsteed
gb_est338:AAV05088	7.00	111.09	2.2e+03	266	AAV05088 vt21n03.x1 Soares_Died	gb_est10:AA161938	7.00	110.44	2.3e+03	289	AAV161938 nm40601.r1 Batsteed
gb_est10:AA129523	7.00	111.07	2.2e+03	267	AA129523 zn79101.s1 Stratagene	gb_est34:AAV136624	7.00	110.44	2.3e+03	291	AAV136624 AV136624 Mus muscul

gb_ests34:AM0768951	+	7.00	110.44	2.3e+03	291	AA168921	Mus musculus	gb_ests25:AA1765865	-	7.00	109.91	2.5e+03	313	AA126865	CG11f09.x1	Soares_N
gb_ests38:AM078455	+	7.00	110.44	2.3e+03	291	AA087455	XB98h08.x1	gb_ests33:AA1775496	-	7.00	109.91	2.5e+03	313	AA177546	EST256396	tomato1.re
gb_ests11:AA210478	+	7.00	110.39	2.4e+03	293	AA210478	mu30h02.x1	gb_ests24:AA1213186	-	7.00	109.89	2.5e+03	314	AA1213186	ye401a1.f1	Aspergill
gb_ests23:AA1390457	+	7.00	110.39	2.4e+03	293	AA1390457	qc22f06.x1	gb_ests37:AA0202112	-	7.00	109.89	2.5e+03	314	AA0202112	ye444504.y1	NCI_CGAP
gb_ests29:AA1559435	+	7.00	110.39	2.4e+03	293	AA1559435	tg33g05.x1	gb_gss3:B38522	-	7.00	109.89	2.5e+03	314	B38522	HS-1047-B2-D9	MF.ab1
gb_ests31:AA0719660	+	7.00	110.39	2.4e+03	293	AA0719660	tg09d11.x1	gb_ests27:AA1603933	-	7.00	109.87	2.5e+03	315	AA1603933	ye481601.y1	Gm-cl004
gb_ests21:AA0737274	+	7.00	110.39	2.4e+03	293	AA0737274	AV037274	gb_ests4:DB05555	-	7.00	109.84	2.5e+03	316	DB05555	HUM117029A	Clontech
gb_ests25:AA1285876	+	7.00	110.37	2.4e+03	294	AA1285876	qh96f07.x1	gb_ests16:AA5633255	-	7.00	109.84	2.5e+03	316	AA563255	V174E01.f1	Knovles
gb_ests37:AA1969407	+	7.00	110.34	2.4e+03	295	AA1969407	wt43c01.x1	gb_ests28:AA176704	-	7.00	109.84	2.5e+03	316	AA176704	tm18c07.x1	NCI_CGAP
gb_ests1:AA19713	+	7.00	110.34	2.4e+03	295	AA19713	924F Heart Homo sapiens	gb_ests3:AA038829	-	7.00	109.84	2.5e+03	316	AA038829	EST280785	tomato1
gb_ests35:AA0408567	+	7.00	110.34	2.4e+03	295	AA0408567	wk2p44310814.x1	gb_ests44:AAV927258	-	7.00	109.84	2.5e+03	316	AAV927258	AV397258	Chlamydom
gb_ests37:AA1970886	+	7.00	110.34	2.4e+03	295	AA1970886	wk2p44310814.x1	gb_ests33:AA024833	-	7.00	109.82	2.5e+03	317	AA024833	wk69e12.f1	NCI_CGAP
gb_gss9:AA054556	+	7.00	110.34	2.4e+03	295	AA054556	CIT-HSP	gb_ests39:AA0100890	-	7.00	109.82	2.5e+03	317	AA0100890	sd62501.y1	Gm-cl008
gb_ests28:AA14744284	+	7.00	110.32	2.4e+03	296	AA14744284	tg8e12.x1	gb_ests4:DB03001	-	7.00	109.80	2.5e+03	318	DB03001	HUM44781B	Clontech
gb_ests29:AA1577939	+	7.00	110.29	2.4e+03	297	AA1577939	UT-R-GO-up-a-11-0-UT-R	gb_ests46:AAW400999	-	7.00	109.80	2.5e+03	318	AAW400999	x239g02.x1	NCI_CGAP
gb_ests16:AA561955	+	7.00	110.27	2.4e+03	298	AA561955	v140b1.f1	gb_gss4:AA0715671	-	7.00	109.80	2.5e+03	318	AA0715671	HS_5429_B1.F03.SP6	
gb_ests127:AA158707	+	7.00	110.27	2.4e+03	298	AA158707	lk13g07.x1	gb_ests4:DB02951	-	7.00	109.77	2.5e+03	319	DB02951	HUM442H02B	Clontech
gb_ests32:AA067109	+	7.00	110.27	2.4e+03	298	AA067109	AV067109	gb_ests24:AA192076	-	7.00	109.77	2.5e+03	319	AA192076	gc95b11.f1	Soares_P
gb_ests29:AA1576907	+	7.00	110.24	2.4e+03	299	AA1576907	UT-R-ABO-vt-f-08-0-UT-R	gb_ests38:AA047030	-	7.00	109.77	2.5e+03	319	AA047030	UT-M-BH1-ahr-a-06-0	
gb_ests33:AA096552	+	7.00	110.24	2.4e+03	299	AA096552	AV096552	gb_ests46:AA134668	-	7.00	109.77	2.5e+03	319	AA134668	10972	MARC
gb_ests33:AA121228	+	7.00	110.24	2.4e+03	299	AA121228	Mus musculus	gb_ests9:CG12129	-	7.00	109.75	2.5e+03	320	CG12129	HUM50002284	Human ad
gb_ests38:AAW040511	+	7.00	110.22	2.4e+03	300	AAW040511	EST82616	gb_ests19:AA0807116	-	7.00	109.75	2.5e+03	320	AA0807116	OC31105.f1	NCI_CGAP
gb_ests42:AAV303874	+	7.00	110.22	2.4e+03	300	AAV303874	AV303874	gb_ests21:AA081139	-	7.00	109.75	2.5e+03	320	AA081139	v557f10.f1	Stratagene
gb_ests6:AA1051	+	7.00	110.22	2.4e+03	300	AA1051	TESTR2y26908.x1	gb_ests29:AA155418	-	7.00	109.75	2.5e+03	320	AA155418	UT-R-C2P-gp-f-05-0	
gb_ests12:AA283046	+	7.00	110.22	2.4e+03	300	AA283046	ct16e11.f1	gb_gss3:B39382	-	7.00	109.75	2.5e+03	320	B39382	HS-1049-B1-G05-MF.ab1	
gb_ests16:CG5714	+	7.00	110.22	2.4e+03	300	CG5714	CG5714	gb_ests15:AA507121	-	7.00	109.73	2.5e+03	321	AA507121	nm41f1.f1	NCI_CGAP
gb_ests16:CG0135	+	7.00	110.22	2.4e+03	300	CG0135	CG0135	gb_ests30:AA143434	-	7.00	109.71	2.5e+03	322	AA143434	V443d08.x1	Soares_LC
gb_ests35:AA156078	+	7.00	110.22	2.4e+03	300	AA156078	AV156078	gb_ests20:AA078964	-	7.00	109.68	2.5e+03	323	AA078964	CJ26C09.f1	NCI_CGAP
gb_ests38:AAW040511	+	7.00	110.22	2.4e+03	300	AAW040511	EST82616	gb_ests2:AA6427	-	7.00	109.66	2.5e+03	324	AA6427	yd87C02.f1	Soares_fet
gb_ests39:AAV204591	+	7.00	110.22	2.4e+03	300	AAV204591	AV204591	gb_ests32:AA122863	-	7.00	109.66	2.5e+03	324	AA122863	fc288a01.f1	zebrafis
gb_ests45:AA252300	+	7.00	110.22	2.4e+03	301	AA252300	UT-R-B12-egy-e-0-0-UT-R	gb_ests9:AA095150	-	7.00	109.64	2.5e+03	325	AA095150	CG282972	seg.F Human
gb_ests11:AA214667	+	7.00	110.20	2.4e+03	301	AA214667	EST72890	gb_ests23:AA155048	-	7.00	109.64	2.5e+03	325	AA155048	ud58010.f1	Soares_N
gb_ests24:AA123886	+	7.00	110.20	2.4e+03	301	AA123886	qh8a06.x1	gb_ests29:AA108165	-	7.00	109.64	2.5e+03	325	AA108165	xv32c05.f1	Stratagene
gb_ests38:AAW040511	+	7.00	110.20	2.4e+03	301	AAW040511	EST82616	gb_ests46:AAW68597	-	7.00	109.64	2.5e+03	325	AAW68597	CV4-HT0104-0911	99-0
gb_gss9:AA002361	+	7.00	110.20	2.4e+03	301	AA002361	1(2)07054	gb_ests20:AA823138	-	7.00	109.62	2.5e+03	326	AA823138	vw41b05.f1	Soares_m
gb_ests13:AAV013139	+	7.00	110.17	2.4e+03	302	AAV013139	EST35258	gb_ests18:AA133100	-	7.00	109.62	2.5e+03	326	AA133100	AV093306	Mus muscul
gb_ests33:AAV011688	+	7.00	110.17	2.4e+03	302	AAV011688	AV011688	gb_ests3:AA093306	-	7.00	109.62	2.5e+03	326	AA093306	AV093306	Mus muscul
gb_ests33:AAW09858	+	7.00	110.17	2.4e+03	302	AAW09858	ws88d03.x1	gb_ests13:AA131300	-	7.00	109.59	2.5e+03	327	AA131300	nv68e05.f1	NCI_CGAP
gb_ests40:AA155925	+	7.00	110.17	2.4e+03	302	AA155925	ga22b01.y1	gb_ests38:AAW06157	-	7.00	109.59	2.5e+03	327	AAW06157	xc77c09.f1	NCI_CGAP
gb_ests12:AA300788	+	7.00	110.15	2.4e+03	303	AA300788	EST13648	gb_ests24:AA1215283	-	7.00	109.57	2.5e+03	328	AA1215283	te6075	MVAT4
gb_ests19:AA154700	+	7.00	110.15	2.4e+03	303	AA154700	v010g08.f1	gb_ests7:W21151	-	7.00	109.53	2.5e+03	330	W21151	zb56a01.f1	Soares_fet
gb_ests13:AAV04700	+	7.00	110.15	2.4e+03	303	AAV04700	UT-R-EO-ch-a-01-0-UT-R	gb_ests2:AA179933	-	7.00	109.51	2.5e+03	331	AA179933	ca19f06.f1	Barstead
gb_ests26:AAW66408	+	7.00	110.15	2.4e+03	304	AAW66408	UT-R-EO-ch-a-01-0-UT-R	gb_ests2:AA179933	-	7.00	109.48	2.5e+03	332	AA179933	ye44a10.f1	Soares_fet
gb_ests1:AA1713	+	7.00	110.12	2.4e+03	304	AA1713	R1C53488A	gb_ests13:AA107239	-	7.00	109.44	2.5e+03	334	AA107239	ve80e07.f1	Soares_N
gb_ests15:AA050601	+	7.00	110.12	2.4e+03	304	AA050601	v19e0a07.f1	gb_ests5:DB6541	-	7.00	109.42	2.5e+03	335	DB6541	CEK071C5R	yu1 Kohar
gb_ests12:AA293388	+	7.00	110.12	2.4e+03	305	AA293388	zt53d11.f1	gb_ests27:AA131074	-	7.00	109.46	2.5e+03	333	AA131074	mf58e09.f1	Soares_m
gb_ests12:AA293388	+	7.00	110.10	2.4e+03	305	AA293388	zt53d11.f1	gb_ests27:AA131074	-	7.00	109.46	2.5e+03	333	AA131074	mf58e09.f1	Soares_m
gb_ests13:AAV010146	+	7.00	110.10	2.4e+03	305	AAV010146	AV010146	gb_ests29:AA108165	-	7.00	109.46	2.5e+03	333	AA108165	xc77c09.f1	NCI_CGAP
gb_ests30:AAV010146	+	7.00	110.10	2.4e+03	305	AAV010146	AV010146	gb_ests29:AA108165	-	7.00	109.46	2.5e+03	333	AA108165	xc77c09.f1	NCI_CGAP
gb_ests30:AAV010146	+	7.00	110.10	2.4e+03	305	AAV010146	AV010146	gb_ests29:AA108165	-	7.00	109.46	2.5e+03	333	AA108165	xc77c09.f1	NCI_CGAP
gb_ests30:AAV010146	+	7.00	110.10	2.4e+03	305	AAV010146	AV010146	gb_ests29:AA108165	-	7.00	109.46	2.5e+03	333	AA108165	xc77c09.f1	NCI_CGAP
gb_ests30:AAV010146	+	7.00	110.10	2.4e+03	305	AAV010146	AV010146	gb_ests29:AA108165	-	7.00	109.46	2.5e+03	333	AA108165	xc77c09.f1	NCI_CGAP
gb_ests30:AAV010146	+	7.00	110.10	2.4e+03	305	AAV010146	AV010146	gb_ests29:AA108165	-	7.00	109.46	2.5e+03	333	AA108165	xc77c09.f1	NCI_CGAP
gb_ests30:AAV010146	+	7.00	110.10	2.4e+03	305	AAV010146	AV010146	gb_ests29:AA108165	-	7.00	109.46	2.5e+03	333	AA108165	xc77c09.f1	NCI_CGAP
gb_ests30:AAV010146	+	7.00	110.10	2.4e+03	305	AAV010146	AV010146	gb_ests29:AA108165	-	7.00	109.46	2.5e+03	333	AA108165	xc77c09.f1	NCI_CGAP
gb_ests30:AAV010146	+	7.00	110.10	2.4e+03	305	AAV010146	AV010146	gb_ests29:AA108165	-	7.00	109.46	2.5e+03	333	AA108165	xc77c09.f1	NCI_CGAP
gb_ests30:AAV010146	+	7.00	110.10	2.4e+03	305	AAV010146	AV010146	gb_ests29:AA108165	-	7.00	109.46	2.5e+03	333	AA108165	xc77c09.f1	NCI_CGAP
gb_ests30:AAV010146	+	7.00	110.10	2.4e+03	305	AAV010146	AV010146	gb_ests29:AA108165	-	7.00	109.46	2.5e+03	333	AA108165	xc77c09.f1	NCI_CGAP
gb_ests30:AAV010146	+	7.00	110.10	2.4e+03	305	AAV010146	AV010146	gb_ests29:AA108165	-	7.00	109.46	2.5e+03	333	AA108165	xc77c09.f1	NCI_CGAP
gb_ests30:AAV010146	+	7.00	110.10	2.4e+03	305	AAV010146	AV010146	gb_ests29:AA108165	-	7.00	109.46	2.5e+03	333	AA108165	xc77c09.f1	NCI_CGAP
gb_ests30:AAV010146	+	7.00	110.10	2.4e+03	305	AAV010146	AV010146	gb_ests29:AA108165	-	7.00	109.46	2.5e+03	333	AA108165	xc77c09.f1	NCI_CGAP
gb_ests30:AAV010146	+	7.00	110.10	2.4e+03	305	AAV010146	AV010146	gb_ests29:AA108165	-	7.00	109.46	2.5e+03	333	AA108165	xc77c09.f1	NCI_CGAP
gb_ests30:AAV010146	+	7.00	110.10	2.4e+03	305	AAV010146	AV010146	gb_ests29:AA108165	-	7.00	109.46	2.5e+03	333	AA108165	xc77c09.f1	NCI_CGAP
gb_ests30:AAV010146	+	7.00	110.10	2.4e+03	305	AAV010146	AV010146	gb_ests29:AA108165	-	7.00	109.46	2.5e+03	333	AA108165	xc77c09.f1	NCI_CGAP
gb_ests30:AAV010146	+	7.00	110.1													

gb_est11:AA032168	-	7.00	109.27	2.7e+03	342	AM0402168	EST15268	Aorta	endothel	gb_est13:AAV171264	+	7.00	108.84	2.9e+03	363	AV171264	AV171264	Mus	muscle		
gb_est18:AA045313	-	7.00	109.27	2.7e+03	342	AA045313	UT1-N-BH1	akr-e	tl-0-Ut	gb_est16:AA065853	+	7.00	108.84	2.9e+03	363	AA065853	CMS-HNO194	Mus	muscle		
gb_est13:AA039379	-	7.00	109.25	2.7e+03	343	AA039379	EST14537	Fetal	brain	gb_gss5:AO077024	+	7.00	108.84	2.9e+03	363	AO770244	HS-2260_B1-H10-TTC				
gb_est7:RS1512	-	7.00	109.23	2.7e+03	344	RS1452	yg16b01.r1	Soares	infant	gb_est8:AA094648	+	7.00	108.82	2.9e+03	364	AA094648	mj57h01.r1	Soares	n		
gb_est12:AA081062	-	7.00	109.23	2.7e+03	344	W53411	md41c05.r1	Soares	mouse	gb_est25:AI080024	+	7.00	108.82	2.9e+03	364	AI080024	qbs5409.r1	Soares	l		
gb_est11:AA077039	-	7.00	109.23	2.7e+03	344	AA081062	zt01d04.r1	NCT	CGAP	gb_est13:AI019105	+	7.00	108.82	2.9e+03	364	AI019105	as77f09.r1	Barstead	hd		
gb_est35:AI035537	-	7.00	109.23	2.7e+03	344	AA077039	vt1lb05.r1	Barstead	mc	gb_est32:AI082708	-	7.00	108.82	2.9e+03	364	AI082708	w228b10.r1	NCT	CGAP		
gb_est33:AI041077	-	7.00	109.23	2.7e+03	344	AI035837	UT-N-A10	aat-a	06-0-Ut	gb_est45:AA029694	-	7.00	108.82	2.9e+03	364	AA029694	UT-H-BW	aj0-e-02-e			
gb_gss3:BB5779	-	7.00	109.21	2.7e+03	345	AI041077	UT-N-AH0	acp-z	08-0-Ut	gb_gss5:BB9541	+	7.00	108.82	2.9e+03	364	BB5441	CIT-HSP-2021P1	TF	C		
gb_gss13:AA041397	-	7.00	109.21	2.7e+03	345	BB5779	RPC111-21B10	TF	RPC1-11	gb_gss5:BB92708	+	7.00	108.80	2.9e+03	365	H97208	yy93f08.r1	Soares	me		
gb_est28:AI094634	-	7.00	109.18	2.8e+03	346	AA041397	RPC1-11	10N2	TF	gb_est29:AI059574	+	7.00	108.80	2.9e+03	365	AI057584	TYR-90	ut-b-04-0-Ut			
gb_est17:AI096815	-	7.00	109.18	2.8e+03	346	AI094634	sb1a12.r1	Yl	GM-1004	gb_est19:AI019934	+	7.00	108.80	2.9e+03	365	AI019934	ty93a07.r1	NCT	CGAP		
gb_est45:AA0269191	-	7.00	109.18	2.8e+03	346	AI096815	wt90f09.r1	NCT	CGAP	gb_est37:AI045864	+	7.00	108.80	2.9e+03	365	AI045864	ts50a01.r1	NCT	CGAP		
gb_est5:HB190	-	7.00	109.16	2.8e+03	347	AA0269191	wt33b11.r1	NCT	CGAP	gb_gss12:AA031106	+	7.00	108.80	2.9e+03	365	AA031106	CITB1-B1	-7526M24	T11		
gb_est11:CT74652	-	7.00	109.16	2.8e+03	347	CT74652	CT74652	Rice	CDNA	from	im	gb_gss14:AA052815	+	7.00	108.78	2.9e+03	366	AA052815	HS_5182	A2-B02	Sp65
gb_est20:AA0901535	-	7.00	109.16	2.8e+03	347	AA091535	NCSMB13	Mycelial	Neut	gb_est7:W04605	-	7.00	108.78	2.9e+03	366	W04605	z498e09.r1	Soares	fet		
gb_est46:AA065849	-	7.00	109.16	2.8e+03	347	AA065849	CMT-H70194	-081099	-02-0	gb_gss10:AA005853	+	7.00	108.78	2.9e+03	366	AA005853	HS_3057	A2-B02	ME		
gb_gss9:AA020551	-	7.00	109.16	2.8e+03	347	AO020551	CIT-HSP-23101290	TFB	CIT	gb_est11:AA0203114	+	7.00	108.76	2.9e+03	367	AA020733	mve99h08.r1	Glyan	mouse		
gb_gss10:AO102739	-	7.00	109.16	2.8e+03	347	AO102739	HS_-050	B1-G09	MF	CIT	gb_est23:AA0023114	+	7.00	108.76	2.9e+03	367	AA0023114	AU023114	Mouse	unf	
gb_est11:MW9118	-	7.00	109.14	2.8e+03	348	AA012407	EST0166	Subtract	ed	gb_gss3:BB9363	+	7.00	108.76	2.9e+03	367	BB9363	CIT-HSP-2053P1	TF	C		
gb_est13:AA1928337	-	7.00	109.08	2.8e+03	351	AA192837	wo96b02.r1	NCT	CGAP	gb_est17:AA0655187	+	7.00	108.74	2.9e+03	368	AA065517	nub2da04.r1	NCT	CGAP		
gb_est47:AAW51112	-	7.00	109.08	2.8e+03	351	AAW51112	UT-H-B13	-ale-c	03-0-Ut	gb_est12:AA0254363	+	7.00	108.74	2.9e+03	368	AA0254363	AO1024503	Mouse	unf		
gb_gss3:BB0584	-	7.00	109.08	2.8e+03	351	BB0584	CIT-HSP-2013P20	TF	CIT	gb_est46:AA061308	+	7.00	108.74	2.9e+03	368	AA061308	RC3-CIT0254	-280999-			
gb_est47:AAW43388	-	7.00	109.06	2.8e+03	352	AAW43388	EST108218	tomato	mixed	gb_est5:HB5703	+	7.00	108.72	2.9e+03	369	HB5703	yg62f09.r1	Soares	fet		
gb_gss9:AA080654	-	7.00	109.06	2.8e+03	352	AA080654	CIT-HSP-223636	TF	CIT	gb_est26:AA0733955	+	7.00	108.72	2.9e+03	369	AA0733955	EST230643	Normalized			
gb_est11:AA394010	-	7.00	109.04	2.8e+03	353	AA394010	zt68a09.r1	Soares	test	gb_est23:AI0808557	+	7.00	108.70	2.9e+03	370	AI088557	qb12c12.r1	Soares	l		
gb_est22:AI441861	-	7.00	109.04	2.8e+03	353	AI441861	sa5c11.r1	GM	-c1004	gb_est12:AI0334023	+	7.00	108.70	2.9e+03	370	AI034023	EST230711	Normalized			
gb_est8:AA010443	-	7.00	109.02	2.8e+03	354	R14456	yg05c09.r1	Soares	infant	gb_est36:AA0189718	+	7.00	108.70	2.9e+03	370	AA018978	AV187708	Yuj1	Kohara		
gb_est11:AA0630912	-	7.00	109.02	2.8e+03	354	AA010443	zt18a03.r1	Soares	fet	gb_est16:AA0894918	+	7.00	108.68	2.9e+03	371	AA089492	n17h09.r1	NCT	CGAP		
gb_est12:AA0981167	-	7.00	109.02	2.8e+03	354	AA081167	wt56d04.r1	Stratagene		gb_est16:AA0548672	+	7.00	108.68	2.9e+03	371	AA054867	sb17h02.r1	Yl	GM		
gb_est26:AA0859579	-	7.00	109.02	2.8e+03	355	AA0859579	UT-R-ED	-bl	-0-Ut	gb_est28:AI0958958	+	7.00	108.68	2.9e+03	371	AI095895	sb17h02.r1	Yl	GM		
gb_est23:AI0589721	-	7.00	109.00	2.8e+03	355	AI0589721	ts11e05.r1	NCT	CGAP	gb_est37:AI0951893	+	7.00	108.68	2.9e+03	371	AI095189	wt38c12.r1	NCT	CGAP		
gb_est3:RT7115	-	7.00	108.98	2.8e+03	356	RT7115	yt164d09.r1	Soares	placent	gb_est33:AA0216330	+	7.00	108.68	2.9e+03	371	AA021630	61407D12	Y2	614		
gb_est13:AA0503149	-	7.00	108.98	2.8e+03	356	AA0503149	ng20e01.r1	NCT	CGAP	gb_est44:AA0021197	+	7.00	108.68	2.9e+03	371	AA002197	sf07d06.r1	GM	-c102		
gb_est38:AA0933828	-	7.00	108.96	2.8e+03	357	AA0933828	EST187008	tomato	mixed	gb_gss6:AA0847006	-	7.00	108.68	2.9e+03	371	AA084706	IM14JFY	Im21d10	Y1		
gb_est46:AA0294414	-	7.00	108.96	2.8e+03	357	AA0294414	UT-H-B12	-sheg	-01-0-Ut	gb_est8:CG08613	+	7.00	108.66	2.9e+03	372	CG08613	CO8613	Yuj1	Kohara		
gb_est33:AI0771611	-	7.00	108.94	2.8e+03	358	AI0771611	EST152711	tomato	ovary	gb_est11:AA0600631	+	7.00	108.66	2.9e+03	372	AA0600631	vm76g12.r1	Knowles			
gb_est6:NA4675	-	7.00	108.92	2.9e+03	359	NA4675	yy22a08.r1	Soares	melanc	gb_est17:CT10186	+	7.00	108.66	2.9e+03	372	CT10186	C71066	Yuj1	Kohara		
gb_est35:AA0100791	-	7.00	108.92	2.9e+03	359	AA0100791	sd61c08.r1	GM	-c1008	gb_est18:AA030099	+	7.00	108.66	2.9e+03	372	AA030099	nx34f09.r1	NCT	CGAP		
gb_est16:CG39521	-	7.00	108.90	2.9e+03	360	CG39521	CG39521	Yuj1	Kohara	unpud	gb_est22:AI0166573	+	7.00	108.66	2.9e+03	372	AI016573	TENU2389	T	crui	
gb_est15:CA43853	-	7.00	108.90	2.9e+03	360	CG39576	CG39576	Yuj1	Kohara	unpud	gb_est23:AI0456533	+	7.00	108.66	2.9e+03	372	AI045653	sb02c10	Y1	GM	
gb_est11:CA67930	-	7.00	108.90	2.9e+03	360	CA43853	CA43853	Yuj1	Kohara	unpud	gb_est38:AA0215636	+	7.00	108.66	2.9e+03	372	AA0215636	AV215436	Riken	Full1	
gb_est17:CG67930	-	7.00	108.90	2.9e+03	360	CA6097	CA6097	Yuj1	Kohara	unpud	gb_est39:AA0215636	+	7.00	108.66	2.9e+03	372	AA0215636	AV215436	Riken	Full1	
gb_est11:CG6904	-	7.00	108.90	2.9e+03	360	CG67930	CG67930	Yuj1	Kohara	unpud	gb_gss11:FR0032978	+	7.00	108.64	2.9e+03	373	FR0032978	Fugu	tubridis	GSS	
gb_est13:CG6904	-	7.00	108.90	2.9e+03	360	CG6904	CG6904	Yuj1	Kohara	unpud	gb_est29:AI056126	+	7.00	108.64	3.0e+03	373	AI056126	vo05f05	X1	Stratagene	
gb_est36:AA0913329	-	7.00	108.90	2.9e+03	360	AA0913329	AV193329	Yuj1	Kohara	unpud	gb_gss11:AA0192889	+	7.00	108.64	3.0e+03	373	AA019288	HS_3045	B1	B07	
gb_est36:AA0201160	-	7.00	108.90	2.9e+03	360	AV0201160	AV201160	Yuj1	Kohara	unpud	gb_est9:CI021208	+	7.00	108.62	3.0e+03	374	CI02708	CI02708	Yuj1	Kohara	
gb_est33:AA0202124	-	7.00	108.90	2.9e+03	360	AV0202124	AV202124	Yuj1	Kohara	unpud	gb_est11:AA028018	+	7.00	108.62	3.0e+03	374	AA028018	ng51e02	X1	NCT	
gb_est13:AA0376461	-	7.00	108.88	2.9e+03	361	AA0376461	EST142219	Endometrial	h	gb_est17:CG9755	+	7.00	108.62	3.0e+03	374	CG9755	CG9755	Yuj1	Kohara		
gb_est12:AI0170543	-	7.00	108.88	2.9e+03	361	AI0170543	EST1216471	Normalized	h	gb_est17:CG9755	+	7.00	108.62	3.0e+03	374	CG9755	CG9755	Yuj1	Kohara		
gb_est13:AI0163814	-	7.00	108.88	2.9e+03	361	AI0163814	UT-R-V0	-acs	-d1-0-Ut	gb_est24:AI0178563	+	7.00	108.62	3.0e+03	374	AI0178563	EST2222241	Normalized	h		
gb_est46:AA0359613	-	7.00	108.88	2.9e+03	361	AA0359613	46375	MARC	2P1G	Sus	sc	gb_est36:AA0214420	+	7.00	108.62	3.0e+03	374	AA0214420	AV0201420	Yuj1	Kohara
gb_est47:AA0414995	-	7.00	108.88	2.9e+03	361	AA0414995	48842	MARC	1P1G	Sus	sc	gb_est11:TA58287	+	7.00	108.60	3.0e+03	375	TA5828	TA5828	Yl	Stratagene
gb_gss9:AA0071198	-	7.00	108.88	2.9e+03	361	AA0071198	HS_-2174	A2	-D02	MR	CIT	gb_est3:RT37179	+	7.00	108.60	3.0e+03	375	RT37179	yt50f09.r1	Soares	bire
gb_est13:AA0813782	-	7.00	108.86	2.9e+03	362	AA0813782	ad17h05	sl	Soares	test	gb_est9:AA012135	+	7.00	108.60	3.0e+03	375	CG0897	CG0897	Yuj1	Kohara	
gb_est33:AI050355	-	7.00	108.86	2.9e+03	362	AI050355	UT-H-B61	-lig	-a-11-0-Ut	gb_est16:CA47485	+	7.00	108.60	3.0e+03	375	CA47485	CA47485	Yuj1	Kohara		
gb_est13:RT23107	-	7.00	108.84	2.9e+03	363	RT23107	ym27a10	r1	Soares	placent	gb_est20:AA0857270	+	7.00	108.60	3.0e+03	375	AA0857270	oh56a05	X1	Knowles	
gb_est30:AI061448	-	7.00	108.84	2.9e+03	363	AI061448	yg34f05	sl	Barstead	hd	gb_est12:CG92985	+	7.00	108.60	3.0e+03	375	CG9298	CG9298	3	Dicystosellum	
gb_est3:AI0811703	-	7.00	108.84	2.9e+03	363	AI0811703	tw75a03	xl	NCT	CGAP	unpud	gb_est11:AA046507	+	7.00	108.60	3.0e+03	375	AA046507	AA046507	Mus	musculu

gb_est37:AA0281501	-	7.00	108.60	3.0e+03	375	AW017501 614059G10.x1 614 - rdc	gb_gss9:BA3587	-	7.00	108.35	3.1e+03	388	B93587 CTT-HSP-2164L23.TR CI
gb_est37:AA286780	-	7.00	108.58	3.0e+03	375	AW285760 4A3B-AAX-B-02.R Anophe	gb_est15:AA478093	-	7.00	108.33	3.1e+03	389	AA478093 zuz3h11.r1 Soares o
gb_est37:AA274281	-	7.00	108.58	3.0e+03	376	AA274281 TGEST7224B08.s1 TGMEx3	gb_est15:AA481073	-	7.00	108.33	3.1e+03	389	AA481073 ab19g03.r1 Stratage
gb_est37:AA353675	-	7.00	108.58	3.0e+03	376	AA353675 EST61856 Activated T-C	gb_est12:AA860727	-	7.00	108.33	3.1e+03	389	AA860727 aJ83d03.y1 Soares_P
gb_est37:CA68785	-	7.00	108.58	3.0e+03	376	CA68785 C46785 Yuj1 Kohara unpu	gb_est20:AA050441	-	7.00	108.33	3.1e+03	389	AA050441 u0u6d05.x1 Soares_N
gb_est37:AA132569	-	7.00	108.58	3.0e+03	376	AA132569 UT-R-C2P-ob-c-01-0-UT	gb_gss7:AA096845	-	7.00	108.33	3.1e+03	389	AA096845 LER2C527R LERG Arab
gb_est37:AA1424563	-	7.00	108.58	3.0e+03	376	AA1424563 t932c02.x1 NCI CGAP_P	gb_gss9:AA001045	-	7.00	108.33	3.1e+03	389	AA001045 HS-2177.AL_MR.G06 C
gb_est37:AA661828	-	7.00	108.58	3.0e+03	376	AA661828 m31n10.x1 Barstead mc	gb_est7:TE6229	-	7.00	108.31	3.1e+03	390	TE6229 yC78C05.s1 Soares Inf
gb_gss3:BA5555	-	7.00	108.58	3.0e+03	376	BA5555 CIT-HSP-342B12.TVC CIT-H	gb_est18:HA1086	-	7.00	108.31	3.1e+03	390	HA1086 EST104780 Rat PC-12 C
gb_est30:AA8621738	-	7.00	108.56	3.0e+03	377	AA8621738 VP7111.s1 Knowles Sol	gb_gss3:BA8194	-	7.00	108.31	3.1e+03	390	BA8194 CIT-HSP-2021P15.TFB C
gb_est30:AA655463	-	7.00	108.56	3.0e+03	377	AA655463 t930f07.x1 NCI CGAP_P	gb_est17:AA54519	-	7.00	108.30	3.1e+03	391	AA54519 UT-R-BD-bd-d-01-0-U
gb_gss3:BA191739	-	7.00	108.56	3.0e+03	377	BA191739 AV191739 Yuj1 Kohara u	gb_est18:AA176755	-	7.00	108.30	3.1e+03	391	AA176755 t167h04.x1 NCI CGAP
gb_gss3:BA8439	-	7.00	108.56	3.0e+03	377	BA8439 HS-1047-B1-H11-ME.abi.ct	gb_est18:AA176090	-	7.00	108.30	3.1e+03	391	AA176090 t167h04.x1 NCI CGAP
gb_gss7:AA023494	-	7.00	108.56	3.0e+03	377	AA023494 RPCI-23-29914.TV RPCI-T	gb_est12:AA174893	-	7.00	108.30	3.1e+03	391	AA174893 ct38h02.x1 Barstead
gb_est4:TA76140	-	7.00	108.54	3.0e+03	378	TA76140 10918 Lambda-PRL2 Arabid	gb_est12:AA175363	-	7.00	108.30	3.1e+03	391	AA175363 ESTE836108.y1 Elm
gb_est17:AA655526	-	7.00	108.54	3.0e+03	378	AA655526 v942d10.r1 Stratage	gb_gss10:AA086195	-	7.00	108.30	3.1e+03	391	AA086195 HS-2177.AL_MR.C
gb_est17:C71017	-	7.00	108.54	3.0e+03	378	C71017 C71017 Yuj1 Kohara unpu	gb_est15:AA47635	-	7.00	108.28	3.1e+03	392	AA47635 zuz3h11.s1 Soares o
gb_est17:C71184	-	7.00	108.54	3.0e+03	378	C71184 C71184 Yuj1 Kohara unpu	gb_est12:AA1756121	-	7.00	108.28	3.1e+03	392	AA1756121 ESTE839606.y1 Elm
gb_est19:AA807823	-	7.00	108.54	3.0e+03	378	AA807823 n96506.s1 NCI CGAP_P	gb_est13:AA102571	-	7.00	108.28	3.1e+03	392	AA102571 AV102571 Mus muscu
gb_est31:AA115051	-	7.00	108.54	3.0e+03	378	AA115051 UT-R-AG1-naa-g-06-0-UT	gb_est10:AA148538	-	7.00	108.28	3.1e+03	392	AA148538 ME000411.ER Egg st
gb_est36:AA195313	-	7.00	108.54	3.0e+03	378	AA195313 AV195313 Yuj1 Kohara u	gb_est17:WA98325	-	7.00	108.26	3.1e+03	393	WA98325 me74a09.r1 Soares mou
gb_est36:AA202469	-	7.00	108.54	3.0e+03	378	AA202469 AV202469 Yuj1 Kohara u	gb_est17:AA610357	-	7.00	108.26	3.1e+03	393	AA610357 np96f10.s1 NCI CGAP
gb_gss15:AA0584723	-	7.00	108.52	3.0e+03	379	AA0584723 RPCI-11-440822.TJ RPCI	gb_est16:AA1346947	-	7.00	108.26	3.1e+03	393	AA1346947 qp54b12.x1 NCI CGAP
gb_est14:H33655	-	7.00	108.52	3.0e+03	379	H33655 EST109854 Rat PC-12 cell	gb_est10:AA1651810	-	7.00	108.26	3.1e+03	393	AA1651810 wb55g02.x1 NCI CGAP
gb_est17:CA63937	-	7.00	108.52	3.0e+03	379	CA63937 C66397 Yuj1 Kohara unpu	gb_est14:AA18180793	-	7.00	108.26	3.1e+03	393	AA18180793 tu21b07.x1 NCI CGAP
gb_est20:AA888181	-	7.00	108.52	3.0e+03	379	AA888181 am37b04.s1 Soares_NFL	gb_est17:AA1970741	-	7.00	108.26	3.1e+03	393	AA1970741 wt14c09.x1 NCI CGAP
gb_est21:AA928054	-	7.00	108.52	3.0e+03	379	AA928054 cm20b05.s1 Soares_NFL	gb_est15:AA301292	-	7.00	108.26	3.1e+03	393	AA301292 xs10e03.x1 NCI CGAP
gb_est21:AA967494	-	7.00	108.52	3.0e+03	379	AA967494 ua06f08.r1 Soares_LHM	gb_gss11:AA0206909	-	7.00	108.26	3.1e+03	393	AA0206909 HS-3238.AL_MR.C
gb_est21:AA1097341	-	7.00	108.52	3.0e+03	379	AA1097341 qb35d08.x1 Soares_Preg	gb_est17:R63735	-	7.00	108.24	3.1e+03	394	R63735 y115906.s1 Soares pla
gb_est36:AA2051254	-	7.00	108.52	3.0e+03	379	AA2051254 AV203454 Yuj1 Kohara u	gb_est17:AA609777	-	7.00	108.24	3.1e+03	394	AA609777 t419g1.s1 Soares t
gb_est45:AA250124	-	7.00	108.52	3.0e+03	379	AA250124 um87f09.y1 Sugeno mous	gb_est18:AA1624100	-	7.00	108.24	3.1e+03	394	AA1624100 ts41h02.x1 NCI CGAP
gb_est8:CA0386	-	7.00	108.50	3.0e+03	380	CA0386 C0386 Yuj1 Kohara unpu	gb_est18:AA084572	-	7.00	108.24	3.1e+03	394	AA084572 w26g02.x1 Soares_D
gb_est14:AA447016	-	7.00	108.50	3.0e+03	380	AA447016 zw90e04.r1 Soares_LocA	gb_gss3:BS92385	-	7.00	108.22	3.1e+03	394	BS92385 CTT-HSP-217LM3.TR CI
gb_est17:AA616991	-	7.00	108.50	3.0e+03	380	AA616991 v178d10.r1 Knowles Sol	gb_est17:AA4408969	-	7.00	108.22	3.1e+03	395	AA4408969 EST03461 Mouse 7.5
gb_est19:AA63249	-	7.00	108.48	3.0e+03	381	AA63249 EST103940 S1. cervetislae	gb_est17:AA448881	-	7.00	108.22	3.1e+03	395	AA448881 ES03461 Mouse 7.5
gb_est17:TA73447	-	7.00	108.48	3.0e+03	381	TA73447 TA73447 Yuj1 Kohara u	gb_est18:AA121458	-	7.00	108.22	3.1e+03	395	AA121458 DKE9762177.F1 762
gb_est15:AA487213	-	7.00	108.48	3.0e+03	381	AA487213 ab19f04.s1 Stratage	gb_est18:AA121458	-	7.00	108.22	3.1e+03	395	AA121458 DKE9762177.F1 762
gb_est28:AA1496235	-	7.00	108.48	3.0e+03	381	AA1496235 sb01f02.y1 GM-10004.G1	gb_est14:AA215694	-	7.00	108.22	3.1e+03	395	AA215694 xep9a04.x1 NCI CGAP
gb_est31:AA1696698	-	7.00	108.48	3.0e+03	381	AA1696698 w056f02.x1 NCI CGAP_P	gb_est14:AA215694	-	7.00	108.22	3.1e+03	395	AA215694 xep9a04.x1 NCI CGAP
gb_est33:AA1764996	-	7.00	108.48	3.0e+03	381	AA1764996 wh57d09.x1 NCI CGAP_K4	gb_gss5:AA0774123	-	7.00	108.22	3.1e+03	395	AA0774123 HS-2050.BD.AL17C
gb_est35:AA1841004	-	7.00	108.48	3.0e+03	381	AA1841004 UT-M-AMO-ado-e-05-0-UT	gb_est14:AA416673	-	7.00	108.20	3.1e+03	396	AA416673 zuz3h11.r1 Soares_N
gb_est5:AA4469	-	7.00	108.47	3.0e+03	382	AA446955 H0MESTR2H1 Human Thymus N	gb_est13:AA1129499	-	7.00	108.20	3.1e+03	396	AA1129499 v02f10.r1 Stratage
gb_est18:AA679585	-	7.00	108.47	3.0e+03	382	AA679585 t419g05.s1 Soares_fetea	gb_est13:AA1129499	-	7.00	108.20	3.1e+03	396	AA1129499 v02f10.r1 Stratage
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P
gb_est27:AA447694	-	7.00	108.47	3.0e+03	382	AA447694 m731c12.x1 Soares_mous	gb_est15:AA1278423	-	7.00	108.20	3.1e+03	396	AA1278423 gm33b02.x1 Soares_P

gb_est117:AA149243	+	7	108.11	3.2e+03	401	AA78923z	nx96d07.s1	NCT_CGAP	GB	gb_gss12:AA067383	+	7	100	107.88	3.3e+03	414	AQ267383	RPC111-69c1.TK	RPC			
gb_est122:AA103845	+	7	108.11	3.2e+03	401	AA103825	ox93h11.x1	Soares_sense	GB	gb_est15:739840	+	7	100	107.86	3.3e+03	415	T39840	EST104506.S	cerevis			
gb_est27:AA148751	+	7	108.11	3.2e+03	401	AA148751	ms32a06.x1	Stratagene	GB	gb_est15:763340	+	7	100	107.86	3.3e+03	415	H63440	YEST104506.S	Soares	fe		
gb_est37:AA189733	+	7	108.11	3.2e+03	401	AA189733	wu21e04.x1	Soares_Dieck	GB	gb_est17:AA632025	+	7	100	107.86	3.3e+03	415	AA632025	np62e01.s1	NCT_CGAP			
gb_est37:AA1904037	+	7	108.11	3.2e+03	401	AA19940737	7014967481	X. thaliana	GB	gb_est17:AA134897	+	7	100	107.86	3.3e+03	415	AA134897	EST731459	Normaliz	ed		
gb_gss17:AA010481	+	7	108.11	3.2e+03	401	AA010481	GSS7C0517	Trypanosome	GB	gb_est32:AA174035	+	7	100	107.86	3.3e+03	415	AA174035	wc32h09.x1	NCT_CGAP			
gb_est14:AA151525	+	7	108.09	3.2e+03	402	AA151525	vd16f03.s1	Knawles_Sol	GB	gb_est34:AA1009391	+	7	100	107.86	3.3e+03	415	AA1009391	wf70a02.x1	Soares	fe		
gb_est14:AA178924	+	7	108.09	3.2e+03	402	AA178924	w12f106.r1	Soares_thym	GB	gb_est38:AA1003653	+	7	100	107.86	3.3e+03	415	AA1003653	AA003653	Lambda	Z11		
gb_est31:AA169588	+	7	108.09	3.2e+03	402	AA169588	ts87c01.x1	NCT_CGAP	GB	gb_gss10:AA047328	+	7	100	107.86	3.3e+03	415	AA047328	32_F22R	PNU01	Normm		
gb_gss16:AA069091	+	7	108.09	3.2e+03	402	AA069091	HS_5414.A2	ELI_Sp6E	RR	gb_est14:AA047328	+	7	100	107.85	3.3e+03	416	AA047328	32_F22R	PNU01	Normm		
gb_est8:AA034316	+	7	108.08	3.2e+03	403	AA034316	z102q02.s1	Soares_fed	GB	gb_est13:AA187278	+	7	100	107.85	3.3e+03	416	AA187278	nl405a1.r1	Aspergyl			
gb_est9:CI18445	-	7	108.08	3.2e+03	403	CI18445	CI18446	Human	placenta	cd	gb_est23:AA113174	+	7	100	107.85	3.3e+03	416	AA113174	UI-R-CCP	hp-d	05-0	
gb_est16:AA565531	+	7	108.08	3.2e+03	403	AA565531	nx42d09.s1	NCT_CGAP	GB	gb_est38:AA0334771	+	7	100	107.85	3.3e+03	416	AA0334771	EST728807	tomato	c		
gb_est16:AA186624	+	7	108.08	3.2e+03	403	AA186624	UI-R-A0-ba-a-04-0-UI-S	Aspergillus	GB	gb_gss3:AB003540	+	7	100	107.85	3.3e+03	416	AB003540	CIT-HSP	-2050D11	TF	C	
gb_est126:AA191192	+	7	108.08	3.2e+03	403	AA1891914	UI-R-A0-ba-a-04-0-UI-S	Aspergillus	GB	gb_gss9:AA192429	+	7	100	107.85	3.3e+03	416	AA192429	MA001352	CBF	Soares	fe	
gb_est45:AA336576	+	7	108.08	3.2e+03	403	AA263576	xm4ba02.x1	NCT_CGAP	GB	gb_est7:AA005933	+	7	100	107.83	3.3e+03	417	AA005933	CIT-HSP	-1326	86c15	TR	
gb_est44:AA447124	+	7	108.06	3.2e+03	404	AA447124	32_F122.A1	GC4_MRC	CIT	gb_est9:AA008381	+	7	100	107.83	3.3e+03	417	AA008381	zh15e05.s1	Soares	fe		
gb_gss5:AA081366	+	7	108.06	3.2e+03	404	AA081366	HS_5315.B2	A03_Sp6E	RR	gb_est12:AA186852	+	7	100	107.81	3.3e+03	418	AA186852	z555g05.s1	NCT_CGAP			
gb_est17:AA050817	+	7	108.06	3.2e+03	404	AA050817	30591	Lambda	-PR12	Arab	gb_est23:AA1123074	+	7	100	107.81	3.3e+03	418	AA1123074	ok37e02.s1	Soares	fe	
gb_est40:AA143766	+	7	108.04	3.2e+03	405	AA143766	EST294062	Normaliz		gb_est24:AA181432	+	7	100	107.81	3.3e+03	418	AA181432	tf75e02.x1	NCT_CGAP			
gb_gss11:AA020297	+	7	108.04	3.2e+03	405	AA020297	RPCC11-45pE	TV	RPCC1-11	gb_est35:AA185817	+	7	100	107.81	3.3e+03	418	AA185817	w143f03.x1	NCT_CGAP			
gb_est5:HH3068	-	7	108.04	3.2e+03	405	HH3068	RPCC11-45pE	TV	RPCC1-11	gb_gss10:AA0142844	+	7	100	107.81	3.3e+03	418	AA0142844	HS_2244_A2	FO1	ME		
gb_est7:W92795	+	7	108.02	3.2e+03	406	W92795	zh80e02.s1	Soares_fetal		gb_gss11:AA0239746	+	7	100	107.81	3.3e+03	418	AA0239746	CIT-HSP	-2286G23	TF	C	
gb_est20:AA58037166	+	7	108.02	3.2e+03	406	AA58037166	od09a01.s1	NCT_CGAP	GB	gb_est16:W57840	+	7	100	107.79	3.3e+03	419	W57840	uy88d10.r1	Soares	fe		
gb_est36:AA1890204	+	7	108.02	3.2e+03	406	AA1890204	wm79g11.x1	NCT_CGAP	GB	gb_est26:AA101679	+	7	100	107.79	3.3e+03	419	AA101679	lh24f10.x1	NCT_CGAP			
gb_gss10:AAQ167524	+	7	108.02	3.2e+03	406	AAQ167524	HS_3164_B1	H10_MRC	CIT	gb_est27:AA1944203	+	7	100	107.79	3.3e+03	419	AA1944203	sa29d04.y1	Gm-cl010			
gb_est3:R22706	+	7	108.00	3.2e+03	407	R22706	yh64b09.r1	Soares	Placed	gb_est45:AA097177	+	7	100	107.79	3.3e+03	419	AA097177	UI-R-BM0	-d1d-c-06-1			
gb_est5:HH37336	+	7	108.00	3.2e+03	407	N37336	18563	Lambda	-PR12	Arabid	gb_est45:AA0305827	+	7	100	107.79	3.3e+03	419	AA0305827	SMO33KMAA6G09SK	O		
gb_est37:AA1938619	+	7	108.00	3.2e+03	407	AA1938619	sb56c08.y1	Gm-cl018	GR	gb_gss11:AA0234820	+	7	100	107.79	3.3e+03	419	AA0234820	HS_2052_B2	ELI	17		
gb_gss13:AA0409126	+	7	108.00	3.2e+03	407	AA0409126	HS_5059_A2	H10_Sp6E	RR	gb_gss16:AA0648033	+	7	100	107.79	3.3e+03	419	AA0648033	RPCC193-Dpnt1	-26111			
gb_est25:AA1820582	+	7	107.99	3.2e+03	408	AA1803962	wk31c07.x1	NCT_CGAP	GB	gb_est17:CI08137	+	7	100	107.78	3.3e+03	420	CI08137	C80137	Mouse	3.5-dpc	C	
gb_est35:AA1828528	+	7	107.99	3.2e+03	408	AA1828528	wk30e10.x1	NCT_CGAP	GB	gb_est12:AA10496772	+	7	100	107.78	3.3e+03	420	AA1049672	an33a12.x1	Gesler			
gb_est36:AA195426	+	7	107.99	3.2e+03	408	AA1894326	EST783969	tomato	callu	gb_est23:AA1141366	+	7	100	107.78	3.3e+03	420	AA1141366	uy41d02.s1	Soares	fe		
gb_est37:AA010904	+	7	107.99	3.2e+03	408	AA010904	ST1G11	Plne	TripleX	gb_est24:AA178188	+	7	100	107.78	3.3e+03	420	AA178188	EST721853	Normaliz	ed		
gb_gss3:BB1226	+	7	107.99	3.2e+03	408	BB1223	CIT-HSP	-2167W7	TR	CIT-HS	gb_est26:AA175590	+	7	100	107.78	3.3e+03	420	AA175590	UI-R-E0	cv-a-07-0-0		
gb_gss7:AAQ32798	+	7	107.99	3.2e+03	408	AAQ332798	nbx00051E237	CUGI	RITC	gb_est33:AA176580	+	7	100	107.78	3.3e+03	420	AA176580	EST757680	tomato	r		
gb_gss10:AAQ118146	+	7	107.99	3.2e+03	408	AAQ118146	HS_3003_A1	A02_MRC	CIT	gb_est33:AA175222	+	7	100	107.78	3.3e+03	420	AA175222	UI-M-BH0	-j1c-e-08-1			
gb_gss11:AAQ219142	+	7	107.99	3.2e+03	408	AAQ219142	HS_3244_A1	ELI_MRC	CIT	gb_est10:AA113326	+	7	100	107.78	3.3e+03	420	AA113326	32866	Lambda	-PR12		
gb_gss11:AAQ230918	+	7	107.99	3.2e+03	408	AAQ203918	HS_2028_B1	E05	TR	CIT	gb_est38:AA0039305	+	7	100	107.78	3.3e+03	420	AA0039305	EST781557	tomato	m	
gb_est5:HH7820	+	7	107.97	3.2e+03	409	HH7820	YEST09b03.s1	Soares_fetal		gb_gss1:AA003398	+	7	100	107.78	3.3e+03	420	AA003398	xet79b01.x1	NCT_CGAP			
gb_est14:AA110154	+	7	107.97	3.2e+03	409	AA410154	YEST0107	Mouse	7.5 dpc	gb_est9:AA0068319	+	7	100	107.78	3.3e+03	420	AA0068319	Fugu	tridripes	GSS		
gb_est17:AA119071	+	7	107.97	3.2e+03	409	AA16919071	vo67d06.r1	Soares_sense		gb_est9:AA0068436	+	7	100	107.76	3.3e+03	421	AA0068436	mm35d02.r1	Stratag			
gb_est19:AA0068680	+	7	107.97	3.2e+03	409	AA0068680	UC11h05.s1	NCT_CGAP	GB	gb_est18:AA113326	+	7	100	107.76	3.3e+03	421	AA113326	32866	Lambda	-PR12		
gb_est35:AA1843059	+	7	107.97	3.2e+03	409	AA1843059	UI-M-AK1-aeg-7b-06-0-UI			gb_est12:AA1015501	+	7	100	107.76	3.3e+03	421	AA1015501	324d08	s1	Soares	fe	
gb_est35:AA1851955	+	7	107.97	3.2e+03	409	AA1851955	UI-M-BH0-a1q-e-03-0-UI			gb_est22:AA0101978	+	7	100	107.76	3.3e+03	421	AA0101978	AA019768	Mouse	elg		
gb_est33:AA0071138	+	7	107.95	3.2e+03	410	AA0071138	Rice	CDNA	fric	gb_est24:AA163473	+	7	100	107.76	3.3e+03	421	AA163473	AA024p40	Hybrid	as		
gb_est40:AA156665	+	7	107.95	3.2e+03	410	AA156665	se31h01.y1	Gm-cl015	GR	gb_est35:AA1852602	+	7	100	107.76	3.3e+03	421	AA1852602	UI-M-BH0	-a1u-c-11-			
gb_est40:AA156665	+	7	107.95	3.2e+03	410	AA156665	se31h01.y1	Gm-cl015	GR	gb_est47:AAW20481	+	7	100	107.76	3.3e+03	421	AAW20481	tl782b01.x1	zebrifish			
gb_gss10:AA0084060	+	7	107.95	3.2e+03	411	AA0084060	HS_2230_B2	H04_MRC	CIT	gb_gss6:AA0830381	+	7	100	107.76	3.3e+03	421	AA0830381	LMaJFV1	lm35e08	x1		
gb_est12:RI14240	+	7	107.93	3.2e+03	411	RI14240	yf79a05.r1	Soares_infant		gb_est18:AA0300740	+	7	100	107.74	3.3e+03	422	AA0300740	HS_3069_B1	CO9	ME		
gb_est38:AA0935300	+	7	107.93	3.2e+03	411	AA0935300	EST786660	tomato	mixed	gb_est7:W66440	+	7	100	107.74	3.3e+03	422	W66440	TSESTY94d07	rl	TMEM		
gb_est38:AA0935300	+	7	107.93	3.2e+03	411	AA0935300	EST786660	tomato	mixed	gb_est7:W73951	+	7	100	107.74	3.3e+03	422	W73951	td78f1f1	s1	Soares	fe	
gb_est38:AA0935300	+	7	107.93	3.2e+03	411	AA0935300	EST786660	tomato	mixed	gb_est10:AA171561	+	7	100	107.74	3.3e+03	422	AA171561	zp32f07.r1	Stratag			
gb_est38:AA0935300	+	7	107.93	3.2e+03	411	AA0935300	EST786660	tomato	mixed	gb_est10:AA171561	+	7	100	107.74	3.3e+03	422	AA171561	zp32f07.r1	Stratag			
gb_est38:AA0935300	+	7	107.93	3.2e+03	411	AA0935300	EST786660	tomato	mixed	gb_est10:AA171561	+	7	100	107.74	3.3e+03	422	AA171561	zp32f07.r1	Stratag			
gb_est38:AA0935300	+	7	107.93	3.2e+03	411	AA0935300	EST786660	tomato	mixed	gb_est10:AA171561	+	7	100	107.74	3.3e+03	422	AA171561	zp32f07.r1	Stratag			
gb_est38:AA0935300	+	7	107.93	3.2e+03	411	AA0935300	EST786660	tomato	mixed	gb_est10:AA171561	+	7	100	107.74	3.3e+03	422	AA171561	zp32f07.r1	Stratag			
gb_est38:AA0935300	+	7	107.93	3.2e+03	411	AA0935300	EST786660	tomato	mixed	gb_est10:AA171561	+	7	100	107.74	3.3e+03	422	AA171561	zp32f07.r1	Stratag			
gb_est38:AA0935300	+	7	107.93	3.2e+03	411	AA0935300	EST786660	tomato	mixed	gb_est10:AA171561	+	7	100	107.74	3.3e+03							

```

gb_est18:AA690372 + 7.00 107.69 3.3e+03 425 1 AA690372 v131c11.r1 Barstead md
gb_est26:AA819198 + 7.00 107.69 3.3e+03 425 1 AA819198 ur-R-A0-al-e02-0-UI-S
gb_gss4:AG0715429 + 7.00 107.69 3.3e+03 425 1 AG0715429 HS-3460.A1-G07.SP6E.RF
gb_est12:Y94344 + 7.00 107.67 3.3e+03 426 1 T94344 y31h10.s1 StrataGene 1u
gb_est3:R63015 + 7.00 107.67 3.3e+03 426 1 R63015 y113c05.r1 Soares p1cent
gb_est26:AA899942 + 7.00 107.67 3.3e+03 426 1 AA899842 ur-R-07-cg-9-02-0-UI-S
gb_est36:AI923693 + 7.00 107.67 3.3e+03 426 1 AI923693 w59e07.x1 NCI CGAP_Lu
gb_est37:AI971834 + 7.00 107.67 3.3e+03 426 1 AI971834 w28h08.x1 NCI CGAP_OY
gb_gss13:AA0406207 + 7.00 107.67 3.3e+03 426 1 AA0406207 HS-3047_B2-G02.SP6E.RF
gb_est29:AI592914 + 7.00 107.66 3.4e+03 427 1 AI592914 vq16b12.y1 Barstead st
gb_est34:AI792950 - 7.00 107.66 3.4e+03 427 1 AI792950 om87f05.y5 NCI CGAP_K1

seq_name: gb_est24:AI213187
seq_documentation_block:
LOCUS AI213187 492 bp mRNA EST 19-OCT-1998
DEFINITION Y801a1.r1 Aspergillus nidulans 24hr asexual developmental and
vegetative cDNA lambda zap library Emericella nidulans cDNA clone
Y801a1 5', mRNA sequence.
ACCESSION AI213187
VERSION AI213187.1 GI:3775129
KEYWORDS EST.
SOURCE Emericella nidulans.
ORGANISM Emericella nidulans.
Eukaryota; Fungi; Ascomycota; Eurotiiales; Trichocomaceae;
Emericella.
REFERENCE 1 (bases 1 to 492)
AUTHORS Kudler,D., Gray,J., Hausner,J., Lai,H., Martin,W., Aramayo,R.,
Priddy,R. and Roe,B.
TITLE An Aspergillus nidulans EST Database
COMMENT Unpublished (1998)
Other ESTs: Y8d01a1.f1
Contact: Bruce A. Roe, University of Oklahoma, broeou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broeou.edu
We anticipate the future release of the cDNA clones to the Fungal
Genetics Stock Center
Seq primer: T3
High quality sequence stop: 431.
Location/Qualifiers
1..492
/organism="Emericella nidulans"
/db_xref="FgSC A26"
/db_xref="taxon:5072"
/clone="Y8d01a1"
/clone_1lb="Aspergillus nidulans 24hr asexual
developmental and vegetative cDNA lambda zap library"
/tissue_type="vegetative mycelia, asexual structures"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
3' end of cDNA cloned into XhoI site of pBluescript"
BASE COUNT 118 a 117 c 132 g 125 t
ORIGIN
alignment_scores:
Quality: 10.00 Length: 10
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
Alignment_block:
US-09-080-127-2 x AI213187 ..
Align seg 1/1 to: AI213187 from: 1 to: 492
255 GYPRGGLYIIEHNSNAPSAPGYSERYLY 254
|||||
86 GGACCTGTATCAATGACGACGCTCTGCG 115

```

```

seq_name: gb_est47:AW398169
seq_documentation_block:
LOCUS AW398169 682 bp mRNA EST 07-FEB-2000
DEFINITION EST297932 L. pennellii trichome, Cornell University Lycopersicon
pennellii cDNA clone cLPT1A10 5', mRNA sequence.
ACCESSION AW398169
VERSION AW398169.1 GI:6916555
KEYWORDS EST.
SOURCE Lycopersicon pennellii.
ORGANISM Lycopersicon pennellii.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 682)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Lakey,J., Holt,I.E.,
Liang,F., Hansen,T.S., Upton,J., Ronning,C.M., Craven,M.B.,
Fuji,C.Y., Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C.,
Martin,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE Generation of ESTs from wild tomato (Lycopersicon pennellii)
trichomes
COMMENT Unpublished (1999)
JOURNAL On Mar 10, 1998 this sequence version replaced gi:2948162.
CONTACT David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
Location/Qualifiers
1..682
/organism="Lycopersicon pennellii"
/db_xref="taxon:28526"
/clone="cLPT1A10"
/clone_1lb="L. pennellii trichome, Cornell University"
/tissue_type="trichome"
/dev_stage="mixed stages"
/lab_host="SOLR"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Leaves of various stages were shaken in liquid
nitrogen, shearing off trichomes. This procedure yielded a
mixture of cells highly enriched for trichomes, with minor
contamination by other types of leaf cells."
BASE COUNT 195 a 129 c 173 g 184 t
ORIGIN
alignment_scores:
Quality: 10.00 Length: 10
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
Alignment_block:
US-09-080-127-2 x AW398169 ..
Align seg 1/1 to: AW398169 from: 1 to: 682
175 MetAlaGlyThrLeuGlyAlaAlaGlnSer 184
|||||
554 ATGGCAGAGACTCTGGAGCTCTCAGAGT 583
seq_name: gb_est34:AI788531
seq_documentation_block:
LOCUS AI788531 278 bp mRNA EST 02-JUL-1999
DEFINITION UK46b04.x1 Sugano mouse kidney mKia mus musculus cDNA clone
IMAGE:1972015 3' similar to gb:M99487 PROSTATE-SPECIFIC MEMBRANE
ANTIGEN (HUMAN);, mRNA sequence.
ACCESSION AI788531
VERSION AI788531.1 GI:5336247

```

```

KEYWORDS      EST.
SOURCE        house mouse.
ORGANISM      Mus musculus.
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS       Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
              1 (bases 1 to 278)
              Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wyllie,T.,
underwood,K., Steptoe,T., Gibbons,M., Page,D., Harvey,N., Schurk,R.,
Ritter,E., Kohm,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
On Dec 20, 1995 this sequence version replaced gi:1134314.
Other ESTs: uk46b04.y1
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:988755
Trace considered overall poor quality
Seq primer: custom primer used
High quality sequence stop: 1.
Location/Qualifiers
    1..278
    /organism="Mus musculus"
    /strain="C57BL"
    /db_xref="taxon:10090"
    /clone="IMAGE:1972015"
    /clone_lib="Sugano mouse kidney mk1a"
    /sex="female"
    /dev_stage="adult"
    /lab_host="DH10B"
    /note="Organ: Kidney; Vector: pME18S-Fl3; Site_1: DraIII
(CACGCTGTG); Site_2: DraIII (CACCAATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGGGCGCCCTTTTCTTTTTTTTTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGGCTCATGTG], digested
and cloned into distinct DraIII sites of the pME18S-Fl3
vector (5' site CACTGTGTG, 3' site CACCAATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGCTCTTAAGCTGC and 3' end
primer CGAAGCTGACGCTGACACA."
BASE COUNT   72 a          64 c          55 g          87 t
ORIGIN
alignment_scores:
    Quality:      9.00           Length:      9
    Ratio:        1.000         Gaps:       0
Percent Similarity: 100.000    Percent Identity: 100.000
Alignment_block:
US-09-080-127-2 x AI788531/rev ..
Align seg 1/1 to reverse of: AI788531 from: 1 to: 278
291 AAtGtGtUphegYLeuLenglyser 239
|||||
|||||
|||||
272 GCAGAGAATTGGCCTTGTTCT 246
seq_name: gb_est3:R35464
seq_documentation block:
LOCUS      R35464          393 bp      mRNA      EST      02-MAY-1995
DEFINITION yN80602.r1 Soares placenta NB2HP Homo sapiens cDNA clone

```

```

IMAGE:136058 5' similar to SP:IP52_ANESU P10280 PROTEASE INHIBITOR
5 II : mRNA sequence.
ACCESSION R35464
VERSION R35464.1 GI:792365
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 393)
AUTHORS Haller,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,D., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The Wash-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1296
High quality sequence stops: 330
Source: IMAGE Consortium, LML
This clone is available royalty-free through LML ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1296 Std Error: 0.00
Seq primer: M13p1
High quality sequence stop: 330.
Location/Qualifiers
1..393
/organism="Homo sapiens"
/db_xref="GDB:542064"
/db_xref="taxon:9606"
/clone="IMAGE:136058"
/clone_id="Soares Placenta Nb2HP"
/sex="Female"
/dev_stage="Placenta obtained at birth (full term)"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: placenta; Vector: pTR3D (Pharmacia) with a
modified polylinker; Site:1: Not I; Site:2: Eco RI; 1st
strand cDNA was primed with a Not I - oligodT primer [5'
AAGTGGAGAGATTGCGCGCGCAGAGATTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pTR73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo. "
BASE COUNT 85 a 103 c 115 g 86 t 4 others
ORIGIN
alignment_scores:
Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-080-127-2 x R35464 ..
Align seg 1/1 to: R35464 from: 1 to: 393
9 LeuLeuSerGlyValLeuAlaGlyArg 17
|||||H|||||H|||||H|||||H
35 CTCCTCTCTGGGGTCTGGCGCGCGCA 61
seq_name: gb-gss14:A0498441
seq_documentation_block:
LOCUS A0498441 452 bp DNA GSS 28-APR-1999
DEFINITION HS-3068_B1.F04.SP6E RPCT-11 Human Male BAC Library Homo sapiens

```


REFERENCE 1 (bases 1 to 465)
AUTHORS Cordonnier-Pratt, M.M., Gingle, A., Pratt, L. and Paterson, A.
TITLE An EST database from Sorghum
JOURNAL Unpublished (1999)
COMMENT On Jan 6, 2000 this sequence version replaced gi:6677422.
Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmp@arches.uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions below Pired quality 16. The threshold for highest quality sequence is 20.
Seq primer: JEN REV
High quality sequence stop: 338
POLYA-No.

FEATURES
source
1..465
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Light Grow 1 (LG1)"
/note="Organ: 10- to 14-day-old light-grown (greenhouse) seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."

BASE COUNT 91 a 155 c 158 g 61 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-080-127-2 x AW287578 ..
Align seg 1/1 to: AW287578 from: 1 to: 465

157 AAlAAlAysAlAysAlAAlAAlAser 165
|||||
114 GCTGCCAAGCGCAAGCGCGCGCTCC 140

seq_name: gb_gss9:AQ079756

seq_documentation_block:
LOCUS AQ079756 467 bp DNA GSS 20-AUG-1998
DEFINITION CIT-HSP-2358E10.TR CIT-HSP Homo sapiens genomic clone 2358E10,
genomic survey sequence.
ACCESSION AQ079756
VERSION AQ079756.1 GI:3440940
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominae; Homo.
1 (bases 1 to 467)
Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Sun, E., Wible, C., Shizuya, H., Simon, M. and
Venter, J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
Unpublished (1998)
Other_GSSs: CIT-HSP-2358E10.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200

Fax: 301 838 0208
Email: mdamas@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M3 Reverse
Class: BAC ends.

FEATURES
source
1..467
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="2358E10"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBeloBAC11; Site_1: HindIII; Site_2: HindIII"

BASE COUNT 143 a 88 c 136 g 100 t
ORIGIN

alignment_scores:
Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-080-127-2 x AQ079756/rev ..
Align seg 1/1 to reverse of: AQ079756 from: 1 to: 467

4 LeuLeuTrpAlaSerLeuLeuSergly 12
|||||
336 CTGCTGTGGGATCCCTGCTGTGCGA 310

seq_name: gb_gss1:CNS00UYM

seq_documentation_block:
LOCUS CNS00UYM 516 bp DNA GSS 28-JUN-1999
DEFINITION Arabidopsis thaliana genome survey sequence SP6 end of BAC 18E17 of
TAMU library from strain Columbia of Arabidopsis thaliana, genomic
survey sequence.
ACCESSION AL091612
VERSION AL091612.1 GI:5292766
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; euroids II; Brassicales; Brassicaceae;
Arabidopsids.
1 (bases 1 to 516)
Salanoubat, M., Choinsne, N., Artiguenave, F., Brottier, P., Wincker, P.,
Samson, D., Saurin, W., Weissenbach, J. and Quetier, F.
Unpublished
2 (bases 1 to 516)
Genoscope.
Direct Submission
Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

FEATURES
source
1..516
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone_lib="TAMU"
/clone_lib="18E17"
/note="end : SP6"
138 a 116 c 115 g 147 t
ORIGIN

Alignment_scores:

Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-080-127-2 x CNS00UYM/rev ..

Align seg 1/1 to reverse of: CNS00UYM from: 1 to: 516

174 SerMetAlaGlyThrLeuGlyAla 182
|||||
468 AGTATGCGAGCGACCTTACGTGACGA 442

seq_name: gb_gss14:A0431848

seq_documentation_block:

LOCUS A0431848 530 bp DNA GSS 31-MAR-1999
DEFINITION HS_5053_A1_E10.SP6E.RPCR-11 Human Male BAC Library Homo sapiens
genomic clone Plate-629 Col-19 Row-I, genomic survey sequence.

ACCESSION A0431848 GI:4542183

KEYWORDS

GSS.

SOURCE

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 530)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome

JOURNALS Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

CONTACT: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallaceu.washington.edu

Clones are derived from the human BAC library RPCR-11. For BAC

library availability, please contact Pieter de Jong

(pieterdejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)

or from Research Genetics (info@resgen.com). BAC end Web Server:

http://www.htsc.washington.edu

Plate: 629 row: I column: 19

Seq primer: SP6

Class: BAC ends

High quality sequence stop: 530.

Location/Qualifiers

1..530

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_11b="RPCR-11 Human Male BAC Library"

/sex="male"

/note="vector: pBAC3.6; Genomic sequence of BAC ends"

BASE COUNT 123 a 138 c 105 g 161 t

FEATURES

Alignment_scores:

Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-080-127-2 x A0431848 ..

Align seg 1/1 to: A0431848 from: 1 to: 530

4 LeuLeuTPAlaSerLeuLeuSer 12

|||||
271 CTGCTGTGGCCCTCCCTGCTGTGGA 297

seq_name: gb_gss5:A0777636

seq_documentation_block:

LOCUS A0777636 558 bp DNA GSS 02-AUG-1999
DEFINITION HS_2252_B2_F04_MR CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=2252 Col=8 Row=L, genomic survey
sequence.

ACCESSION A0777636

VERSION A0777636.1 GI:5680596

KEYWORDS

GSS.

SOURCE

human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 558)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome

JOURNALS Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

CONTACT: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallaceu.washington.edu

Clones may be purchased from Research Genetics (info@resgen.com).

BAC end Web Server: http://www.htsc.washington.edu

Plate: 2252 row: L column: 8

Seq primer: M13 Reverse

Class: BAC ends

High quality sequence stop: 558.

Location/Qualifiers

1..558

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_11b="CIT Approved Human Genomic Sperm Library D"

/sex="male"

/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"

BASE COUNT

153 a 119 c 140 g 144 t 2 others

FEATURES

Alignment_scores:

Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-080-127-2 x A0777636/rev ..

Align seg 1/1 to reverse of: A0777636 from: 1 to: 558

3 SerLeuLeuTPAlaSerLeuLeuSer 11

|||||
311 TCCCTGTGTGGCATCCTCTGTCT 285

seq_name: gb_gss5:AA116793

seq_documentation_block:

LOCUS AA116793 662 bp mRNA EST 13-FEB-1997
DEFINITION mq24h05.r1 Batshead MFLRBI Mus musculus cDNA clone IMAGE:579705 5'
similar to gb:M99487 PROSTATE-SPECIFIC MEMBRANE ANTIGEN (HUMAN)).

mRNA sequence.
 ACCESSION A116793
 VERSION A116793.1 GI:1671805
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 662)
 AUTHORS Maria, M., Hallier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Scheinberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1395417.
 Contact: Maria M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNC. Contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:35433
 Seq primer: -28ml3 rev2 ET from Amerisham
 High quality sequence stop: 270.
 Location/Qualifiers
 1..662
 /organism="Mus musculus"
 /strain="BALB/c"
 /db_xref="taxon:10090"
 /clone_image="579705"
 /clone_lib="Barstead MPLRB1"
 /sex="mixed"
 /tissue_type="kidney"
 /dev_stage="6 weeks"
 /lab_host="DH10B"
 /note="Vector: p7713D-Pac (Pharmacia) with a modified
 polylinker. Site 1: EcoRI; Site 2: NotI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TGTTCGATCTCATAGCGAGCGCGCCCTTTTCTTTTCTTTTCTTTTCTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 [CATGATTCGATCC], digested with Not I and cloned into the
 Not I and Eco RI sites of the modified p7713 vector.
 Library constructed by Bob Barstead."
 BASE COUNT 204 a 129 c 148 g 180 t 1 others
 ORIGIN
 Alignment_scores:
 Quality: 9.00 Length: 9
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 Alignment_block:
 US-09-080-127-2 x A116793 ..
 Align seg 1/1 to: A116793 from: 1 to: 662
 291 AlaGluGluPheGlyLeuLeuGlySer 299
 ||||||||||||||||||||
 70 GCAGAGAGATTGGCCCTCTGCTCT 96
 seq_name: gb_gss12:A0274419
 seq_documentation_block:
 LOCUS A0274419 692 bp DNA GSS 03-NOV-1998
 DEFINITION mgxb0015M12r CUGI Rice Blast Library Pyricularia grisea genomic
 clone mgxb0015M12r, genomic survey sequence.
 ACCESSION A0274419

VERSION A0274419.1 GI:3827734
 KEYWORDS GSS.
 SOURCE Pyricularia grisea.
 ORGANISM Pyricularia grisea
 Eukaryota; Fungi; Ascomycota; anamorphic Ascomycota; Pyricularia.
 REFERENCE 1 (bases 1 to 692)
 AUTHORS Yu, Y., Zhu, H., Boyd, C. A., Gaudette, B., Gayle, A., Kingsbury, R.,
 Phillips, K., Sasnowski, M., Wing, R. A. and Dean, R. A.
 TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
 Genome
 JOURNAL Unpublished (1998)
 COMMENT Contact: Dean RA
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: GGAACAGCATATGACATG
 Class: BAC ends
 High quality sequence stop: 390.
 Location/Qualifiers
 1..692
 /organism="Pyricularia grisea"
 /strain="70-15"
 /db_xref="taxon:89476"
 /clone="mgxb0015M12r"
 /clone_lib="CUGI Rice Blast BAC Library"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /note="Vector: pBACWICH. Site 1: HindIII; Site 2: HindIII;
 Rice blast is one of the most devastating fungal diseases
 of rice world wide. It is a filamentous ascomycete with
 a haploid genome (n=7) of approximately 40 Mbp. Rice
 blast is an important model fungal pathogen for studying
 numerous aspects of the fungal-host interaction. In
 order to facilitate genome wide analysis, a BAC library
 containing 9216 clones with an average insert size of 130
 kbp was constructed. This library represents greater
 than 25X genome coverage. High density colony filters
 are available upon request."
 BASE COUNT 145 a 169 c 210 g 167 t 1 others
 ORIGIN
 Alignment_scores:
 Quality: 9.00 Length: 9
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 Alignment_block:
 US-09-080-127-2 x A0274419/rev ..
 Align seg 1/1 to reverse of: A0274419 from: 1 to: 692
 157 AlaAlaLysAlaLysAlaAlaLaser 165
 ||||||||||||||||||||
 110 GCTGCAAGAGCCAGCGCCGCGTCC 84
 seq_name: gb_est10:A152553
 seq_documentation_block:
 LOCUS A152553 124 bp mRNA EST 11-FEB-1997
 DEFINITION m85f03.r1 Stragene mouse heart (#937316) Mus musculus cDNA clone
 IMAGE:604253 5' similar to TR:G510307 G510307 TRANSLATION
 INITIATIONFACTOR EIF-4GAMMA ;, mRNA sequence.
 ACCESSION A152553
 VERSION A152553.1 GI:1724213
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 124)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
 TITLE The WashU-HHMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT On May 9, 1995 this sequence version replaced g1:802370.
 Contact: Marra M/Mouse EST Project
 WashU-HHMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:369685
 Possible reversed clone: similarity on wrong strand
 Seq primer: -28m13 rev1 ET from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers
 1..124
 /organism="Mus musculus"
 /strain="NIH/Swiss"
 /db_xref="taxon:10090"
 /clone="IMAGE:604253"
 /clone_1lb="Stratagene mouse heart (#937316)"
 /sex="pooled"
 /tissue_type="heart"
 /dev_stage="13 day embryos"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: heart; Vector: pBluescript SK-; Site:1: EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer: 0150 dt. 93 pooled NIH/Swiss 13 day embryo hearts. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGACGACG 3' -3' adaptor sequence: 5' CTCGAGTGTGTGTGTGTGTGT 3'."

BASE COUNT 26 a 34 c 30 g 34 t
 ORIGIN

alignment_scores:
 Quality: 8.00 Length: 8
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-080-127-2 x AA152553 ..
 Align seg 1/1 to: AA152553 from: 1 to: 124

8 SerLeuLeuSerGlyValLeuAla 15
 ||||||||||||||||||||
 77 AGCTGTGCTTAGCGGCTCTTCT 100

seq_name: gb_est19:AA790684

seq_documentation_block:
 LOCUS AA790684 148 bp mRNA EST 06-FEB-1998
 DEFINITION vW18C08.r1 Soares_mammary_gland_NbMWG Mus musculus cDNA clone
 IMAGE:1244174 5', mRNA sequence.
 ACCESSION AA790684
 VERSION AA790684.1 GI:2850804
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 148)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
 TITLE The WashU-HHMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT On Jan 9, 1998 this sequence version replaced g1:937862.
 Contact: Marra M/Mouse EST Project
 WashU-HHMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:657862
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 113.
 Location/Qualifiers
 1..148
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:1244174"
 /clone_1lb="Soares_mammary_gland_NbMWG"
 /sex="male"
 /tissue_type="mammary gland"
 /dev_stage="4 weeks"
 /lab_host="DHI0B"
 /note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site:2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTCCATCTGTAGTGGAGCGCCGCAATGCTTTTGTGTGTGTGTGT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Fatima Bonafide."

BASE COUNT 39 a 44 c 38 g 27 t
 ORIGIN

alignment_scores:
 Quality: 8.00 Length: 8
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-080-127-2 x AA790684/rev ..
 Align seg 1/1 to reverse of: AA790684 from: 1 to: 148

172 AAlaglySerMetAlaglyThrIleu 179
 ||||||||||||||||||||
 95 GCCGCGACATGCGCGGACCTC 72

Thu May 11 10:03:11 2000

us-09-080-127-2.rst

Page 15

